

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : A61K 48/00, C12N 5/00, 15/00		A1	(11) International Publication Number: WO 97/00698
			(43) International Publication Date: 9 January 1997 (09.01.97)
(21) International Application Number: PCT/US96/10602		(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(22) International Filing Date: 20 June 1996 (20.06.96)			
(30) Priority Data: 60/017,814 20 June 1995 (20.06.95) US		Published With international search report.	
(71) Applicant: THE GENERAL HOSPITAL CORPORATION [US/US]; 55 Fruit Street, Boston, MA 02114 (US).			
(72) Inventors: WANDS, Jack, R.; 210 Varick Road, Waban, MA 02168 (US). SCAGLIONI, Pier, Paolo; 6 Sheafe Street #4, Boston, MA 02113 (US). MELEGARI, Margherita; 6 Sheafe Street #4, Boston, MA 02113 (US).			
(74) Agent: CLARK, Paul, T.; Fish & Richardson P.C., 225 Franklin Street, Boston, MA 02110-2804 (US).			
(54) Title: INHIBITION OF HEPATITIS B REPLICATION			
(57) Abstract			
<p>The invention features a method of inhibiting the replication of a naturally-occurring hepadnavirus, e.g., hepatitis B virus (HBV), by introducing into proximity with the hepadnavirus a nucleic acid that encodes a hepadnavirus mutant polypeptide. The polypeptide includes a first amino acid sequence that is substantially identical to a corresponding region of a wild type hepadnavirus core protein, and either lacks a second amino acid sequence of the wild type hepadnavirus core protein, the second sequence including the carboxyterminal three amino acids of the wild type hepadnavirus core protein, and/or is joined by a peptide bond to the aminoterminal amino acid of an amino acid sequence that is substantially identical to a corresponding portion of a wild type hepadnavirus surface protein, the aminoterminal amino acid of the surface protein being joined by a peptide bond to the carboxyterminal amino acid of the core protein sequence.</p>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

- 1 -

INHIBITION OF HEPATITIS B REPLICATION

The invention was funded in part by grants CA-35711 and AA-08169 from the National Institutes of Health. The Government has certain rights to this invention.

Background of the Invention

The invention relates to treating infections of a hepadnavirus, e.g., hepatitis B virus.

10       Hepatitis B virus (HBV) is a member of the hepadnavirus family, a group of enveloped DNA viruses that cause acute and chronic hepatitis. Major clinical consequences of HBV infection include acute liver failure, liver cirrhosis, and primary hepatocellular  
15 carcinoma (HCC). With more than 250 million individuals infected worldwide, effective treatment of chronic HBV infection is a major public health goal (Ganem et al., *Annu. Rev. Biochem.*, 56: 651-693, 1987). Although an effective and inexpensive vaccine is available for  
20 preventing infection, to date there is no effective therapy for treating individuals with persistent infection, nor for reducing the risk of liver disease in infected patients (Maynard et al., *Rev. Infect. Dis.*, 11, S574-S578, 1989); DiBisceglie et al., *Cancer Detection*  
25 *and Prevention*, 14, 291-293, 1989). Current treatments for chronic HBV infection include interferon and other inhibitors of viral DNA synthesis. Since these agents have achieved only limited success, additional antiviral approaches are urgently needed.

30       Hepadnaviruses are composed of a viral envelope, a nucleocapsid which contains a relaxed circular 3.2 kb DNA genome, and a virally encoded reverse transcriptase. Following infection of a cell, virion DNA is delivered to the nucleus where it is converted into a covalently  
35 closed circular DNA (cccDNA), which is in turn

- 2 -

transcribed into several subgenomic and pregenomic mRNAs. The pregenomic RNA is then encapsidated into the viral nucleocapsid, together with the reverse transcriptase enzyme necessary to generate the viral DNA genome (Enders et al., *J. Virol.*, 67, 35-41, 1987). Selective encapsidation of pregenomic RNA depends on both nucleocapsid protein and on viral polymerase (Bartenschlager et al., *J. Virol.*, 64, 5324-5332, 1990; Hirsch et al., *Nature*, 344, 552-555, 1990; Nassal, M., *J. Virol.*, 66, 4107-4116, 1992; Roychoury et al., *J. Virol.*, 65, 3617-3624, 1991) as well as on a cis-acting encapsidation signal located at the 5' end of the pregenomic RNA (Bartenschlager et al., *supra*; Junker-Niepmann et al., *EMBO J.*, 2, 3389-3396, 1990; Pollack et al., *J. Virol.*, 67, 3254-3263, 1993).

The mammalian hepadnavirus 21 kd core protein is a 183-187 (depending on the viral strain) amino acid monomer, 180 of which self assemble into an icosahedral structure within the cytoplasm of infected cells. The core protein has two functional domains. The aminoterminal (amino acids 1 to 139-44) is essential for core assembly. A carboxyterminal arginine-rich region (amino acids 139-183, or 144-187, depending upon the viral strain) binds nucleic acids that are required for positive strand DNA synthesis, and stabilizes core particles for complete assembly of the complex into an enveloped viral particle (Birnbaum et al. *J. Virol.*, 64, 3319-3330, 1990; Yu et al., *J. Virol.*, 65, 2511-2517, 1990; Nassal, M., *supra*).

30

#### Summary of the Invention

The invention is based on Applicants' discovery that altering the carboxyterminus of the hepadnavirus core protein creates a mutant polypeptide that reduces replication of a wild type hepadnavirus, by a dominant

- 3 -

negative mechanism. The inhibitory effect is achieved by deletion of a few carboxyterminal amino acids from the core protein, and/or by joining the core protein to a hepadnavirus surface protein, thereby creating a core-surface fusion polypeptide.

Accordingly, the invention features a method of inhibiting the replication of a naturally-occurring, infectious hepadnavirus. The method involves introducing into the proximity of the hepadnavirus a hepadnavirus mutant polypeptide, or a nucleic acid that encodes such a hepadnavirus mutant polypeptide. The polypeptide includes a first amino acid sequence that is substantially identical to a region of a wild type hepadnavirus core protein, but lacks a second amino acid sequence of the wild type hepadnavirus core protein, wherein the second sequence includes the carboxyterminal three amino acids of the wild-type hepadnavirus core protein and does not exceed 100 amino acids in length. The mutant polypeptide is introduced into the infected cell, or is expressed from the nucleic acid, in the proximity of the naturally-occurring hepadnavirus, so as to be available to inhibit replication of the hepadnavirus.

When the method of inhibiting hepadnavirus replication is targeted against HBV, the carboxyterminal amino acid of the first amino acid sequence can be selected from the group consisting of any of the amino acids between position 81 and position 180 of the sequence shown in Fig. 15 (SEQ ID NO: 12), inclusive; preferably the carboxyterminal amino acid is chosen from the group consisting of the amino acids between position 171 and position 180 of the sequence shown in Fig. 15 (SEQ ID NO: 12), inclusive. A construct exemplified herein ends with a carboxyterminal residue at position 171, so that the mutant core protein includes amino acids

- 4 -

1-171 (Fig. 15 (SEQ ID NO: 12)). In another example, the carboxyterminal amino acid is amino acid 178, so that the mutant core protein includes amino acids 1-178 (Fig. 15 (SEQ ID NO: 12)), corresponding to a five amino acid deletion from the carboxyterminus (see, e.g., the analogous duck hepatitis B virus (DHBV) construct pBK, which is described below). The first amino acid sequence is at least 70 amino acids in length, e.g., 72, 74, 76, 78, or 80 amino acids in length. The aminoterminal amino acid of the first amino acid sequence can be the first amino acid of the corresponding wild type hepadnavirus sequence. Alternatively, nonessential aminoterminal amino acids can be eliminated from the mutant polypeptide, provided that the resulting mutant polypeptide does not lose substantial inhibitory activity as a result, when tested according to the methods described below.

By "lacks a second amino acid sequence" is meant that at least three amino acids from the carboxyterminal end of the core protein have been deleted to make the mutant. Preferably, the deleted sequence includes amino acids 171-183 of the HBV core protein; i.e., the second amino acid sequence includes amino acids 171-183 of the sequence shown in Fig. 15 (SEQ ID NO: 12), inclusive.

In another embodiment of the method of inhibiting hepadnavirus replication, the mutant polypeptide further includes a third amino acid sequence. The third amino acid sequence is substantially identical to a portion of a wild type hepadnavirus surface protein. The aminoterminal amino acid of the third amino acid sequence may be joined by a peptide bond to the carboxyterminal amino acid of the first amino acid sequence so as to create a fusion protein. The third amino acid sequence can be the entire surface protein, or can be a portion thereof, e.g., a portion of at least 4, 8, 20, 30, or 43

- 5 -

amino acids in length. For example, the aminoterminal amino acid of the third amino acid sequence can be selected from the group consisting of the amino acids between position 1 and position 112 of the sequence shown in Fig. 16 (SEQ ID NO: 14), inclusive, preferably the amino acids between position 1 and position 8, inclusive. Preferred aminoterminal amino acids of the third amino acid sequence exemplified herein include, but are not limited to, position 5 or position 8 of Fig. 16 (SEQ ID NO: 14).

The carboxyterminal amino acid of the third amino acid sequence can be selected from a group that includes any of the amino acids between position 51 and position 224 of Fig. 16 (SEQ ID NO: 14), inclusive; e.g., any of the amino acids between position 112 and position 224 of Fig. 16 (SEQ ID NO: 14), inclusive; e.g., the carboxyterminal amino acid may be position 51, position 112, or position 224 of Fig. 16 (SEQ ID NO: 14). Thus, the portion of the surface protein included on the mutant polypeptide preferably includes surface protein residues 1-112, 8-112, or 8-51, all inclusive (Fig. 16; SEQ ID NO: 14).

The use of a core protein for inhibiting viral replication is a species-specific event, so that mutant core proteins inhibit nucleocapsid assembly in the same type of hepadnavirus from which they were derived. Thus, the first amino acid sequence is substantially identical to a region of a wild type hepadnavirus core protein that is derived from the same type of hepadnavirus (e.g., HBV versus DHBV) as the naturally-occurring hepadnavirus targeted for inhibition. In contrast, the third amino acid sequence may be substantially identical to a portion of a wild type hepadnavirus surface protein of any hepadnavirus species, since the surface proteins do not demonstrate species specificity. Thus, when the method

- 6 -

of the invention is used to treat an HBV infection, the mutant polypeptide should include sequences specifically derived from the HBV core protein Fig. 15 (SEQ ID NO: 12), but can include sequences derived from any species of surface protein (e.g., the sequence of Fig. 16 (SEQ ID NO: 14)).

In another embodiment, the invention features a nucleic acid that encodes a mutant hepatitis B virus (HBV) polypeptide, the polypeptide including a first amino acid sequence that is substantially identical to a region of a wild type HBV core protein, and lacking a second amino acid sequence of the wild type HBV core protein. The second sequence includes the carboxyterminal three amino acids of the wild type HBV core protein and does not exceed nine amino acids in length. Thus, the carboxyterminal amino acid of the first amino acid sequence can be at position 174, position 175, position 176, position 177, position 178, position 179, or position 180, all of Fig. 15 (SEQ ID NO: 12).

In another embodiment, the invention features a nucleic acid that encodes a mutant hepadnavirus polypeptide. The polypeptide includes a first amino acid sequence that is substantially identical to a region of a wild type hepadnavirus core protein; lacks a second amino acid sequence of the wild type hepadnavirus core protein which includes at least the carboxyterminal three amino acids of the wild type hepadnavirus core protein; and includes a third amino acid sequence that is substantially identical to a portion, or all, of a wild type hepadnavirus surface protein. The aminoterminal amino acid of the third amino acid sequence may be joined by a peptide bond to the carboxyterminal amino acid of the first amino acid sequence so as to create a fusion protein. The carboxyterminal amino acid of the first

- 7 -

amino acid sequence can be any of the amino acids between position 71 and position 180 of Fig. 15 (SEQ ID NO: 12), inclusive. Preferably, the second amino acid sequence does not exceed 100 amino acids in length.

5       The invention also features polypeptides encoded by any of the various nucleic acids of the invention. A polypeptide of the invention can be included in a therapeutic composition as an active ingredient, along with a pharmaceutically acceptable carrier, or it can be  
10 expressed from the nucleic acid within the infected cell.

---       The invention also features vectors into which are inserted any of the various nucleic acids of the invention. The vector can include any sequence known to those of skill in the art necessary or desirable for  
15 replicating the vector in a eukaryotic cell or for expressing a polypeptide of the invention from the coding sequences thereon. For example, the nucleic acid sequence can be operatively linked to appropriate transcription and/or translation control sequences that  
20 function in a eukaryotic cell. The vector can be any vector suitable for maintaining or making multiple copies of a nucleic acid of the invention, or can be one that is suitable for administering a nucleic acid of the invention to a cell or to a mammal infected with a  
25 hepadnavirus, e.g., to a human patient infected with HBV or to cells removed from the patient for ex vivo gene therapy. Examples of vectors useful in the method of inhibiting a hepadnavirus include, but are not limited to, adenovirus vectors, adeno-associated vectors, and  
30 retroviral vectors. Any of the various vectors of the invention can be included in a therapeutic composition along with a pharmaceutically acceptable carrier.

      In another aspect the invention includes a method of evaluating a candidate polypeptide for its ability to  
35 inhibit the replication of a naturally-occurring

- 8 -

hepadnavirus. The method involves introducing a mutant  
hepadnavirus polypeptide as described above into a medium  
in the presence of the hepadnavirus and determining  
whether hepadnavirus replication is inhibited in the  
5 presence of the polypeptide, compared to in its absence,  
such inhibition being an indication that the polypeptide  
is an inhibitor of hepadnavirus replication. By "medium"  
is meant an environment that is capable of supporting  
viral replication by virtue of its chemical composition.  
10 The medium can be within an organism, e.g., an animal  
model, or can be within an organ removed from an animal.  
The medium can also be an intracellular medium, e.g., in  
a cell culture assay, or a cell-free extract, e.g., a  
cell free replication system. Examples of cells suitable  
15 for a cell culture assay include, but are not limited to,  
Huh-6, Huh-7, HepG2, HepG2 2215, LMH, DC, and HCC cells.  
The polypeptide can be introduced to the medium by  
introducing into the medium a nucleic acid encoding the  
polypeptide, with subsequent expression of the  
20 polypeptide therein.

Another method of inhibiting the replication of a  
naturally-occurring hepadnavirus involves introducing  
into the proximity of the hepadnavirus a hepadnavirus  
mutant polypeptide, or a nucleic acid that encodes a  
25 hepadnavirus mutant polypeptide. The polypeptide  
includes a first amino acid sequence that is  
substantially identical to a region of, or all of, a wild  
type hepadnavirus core protein, and a second amino acid  
sequence which is substantially identical to a portion  
30 of, or all of, a wild type hepadnavirus surface protein.  
The aminoterminal amino acid of the second amino acid  
sequence may be joined by a peptide bond to the  
carboxyterminal amino acid of the first amino acid  
sequence so as to create a fusion protein. The second  
35 amino acid sequence can be the entire surface protein, or

- 9 -

can be a portion thereof. The mutant polypeptide is expressed from the nucleic acid in the proximity of the naturally-occurring hepadnavirus, so as to be available to inhibit replication of the hepadnavirus.

5 In a final aspect, the invention includes a hepadnavirus mutant polypeptide, or a nucleic acid that encodes a hepadnavirus mutant polypeptide. The polypeptide includes a first amino acid sequence that is substantially identical to a region, or all, of a wild  
10 type hepadnavirus core protein, and a second amino acid sequence which is substantially identical to a portion, or all, of a wild type hepadnavirus surface protein. The aminoterminal amino acid of the second amino acid sequence may be joined by a peptide bond to the  
15 carboxyterminal amino acid of the first amino acid sequence so as to create a fusion protein. The second amino acid sequence can be the entire surface protein, or can be a portion thereof.

As used herein, a "hepadnavirus" refers to a  
20 member of the hepadnavirus family of viruses, including, but not limited to, hepatitis B virus and hepatitis delta virus (Wang et al., Nature, 323:508-13, 1986). Although treatment of HBV is an important feature of the method of invention due to the incidence of HBV-related human  
25 disease, the methods described herein also apply to other species of hepadnaviruses. Examples of hepadnaviruses within the scope of the invention include, but are not limited to, hepadnaviruses infecting various human organs, including liver cells, exocrine and endocrine  
30 cells, tubular epithelium of the kidney, spleen cells, leukocytes, lymphocytes, e.g., splenic, peripheral blood, B or T lymphocytes, and cells of the lymph nodes and pancreas (see, e.g., Mason et al., Hepatology, 9:635-645, 1989). The invention also applies to hepadnaviruses  
35 infecting non-human mammalian species, such as

- 10 -

domesticated livestock or household pets. In addition, the invention includes a method of evaluating a candidate mutant polypeptide for its ability to inhibit hepadnavirus replication. For the purposes of conducting  
5 a laboratory screening assay, a variety of hepadnavirus species are useful models. Examples include, but are not limited to, woodchuck hepatitis virus (WHV; Summers et al. Proc. Natl. Acad. Sci. USA, 75:4533-37, 1978), duck hepatitis B virus (DHBV; Mason et al. J. Virol. 36:829-  
10 36, 1978), and squirrel hepatitis virus (e.g., Marion et al. Proc. Natl. Acad. Sci. USA, 77:2941-45, 1980).

Although particular amino acids are referred to below with reference to the sequence of HBV (Figs. 15 and 16; SEQ ID NOs: 11-14), it is understood that the  
15 invention encompasses mutant polypeptides comprising corresponding amino acid segments derived from other hepadnavirus species. One of ordinary skill in the art can easily compare closely-related sequences to locate the analogous amino acid positions in related  
20 hepadnaviruses; the descriptions provided in Examples 2 and 3 illustrate examples of such comparisons.

Where the method of inhibiting hepadnavirus replication is used to treat a hepadnaviral infection in an animal, a "naturally-occurring" hepadnavirus refers to  
25 a form or sequence of the virus as it exists in an animal, e.g., a natural isolate derived from an infected animal. In all other contexts, a "naturally-occurring" hepadnavirus is intended to be synonymous with the sequence known to those skilled in the art as the "wild  
30 type" sequence, e.g., the wild type HBV core and surface protein sequences shown in Figs. 15 and 16 (SEQ ID NOs: 11-14). If an amino acid sequence of a core or surface protein of a hepadnavirus that is derived from a natural isolate differs from the conventionally accepted "wild  
35 type" sequence, it is understood that the sequence of the

- 11 -

natural isolate may be the proper comparison sequence for designing mutant polypeptides of the invention.

"Sequence identity", as used herein, refers to the subunit sequence similarity between two nucleic acid or polypeptide molecules. When a given position in both of the two molecules is occupied by the same nucleotide or amino acid residue, e.g., if a given position in each of two polypeptides is occupied by serine, then they are identical at that position. The identity between two sequences is a direct function of the number of matching or identical positions, e.g., if half (e.g., 5 positions in a polymer 10 subunits in length) of the positions in two polypeptide sequences are identical, then the two sequences are 50% identical; if 90% of the positions, e.g., 9 of 10, are matched, the two sequences share 90% sequence identity. Methods of sequence analysis and alignment for the purpose of comparing the sequence identity of two comparison sequences are well known by those skilled in the art. By "substantially identical" is meant sequences that differ by no more than 10% of the residues, and only by conservative amino acid substitutions such as those shown in Table 1, or non-conservative amino acid substitutions, deletions, or insertions that do not appreciatively diminish the polypeptide's biological activity, e.g., an insertion of amino acids at the junction of the core protein and surface protein sequences that has no appreciative effect on biological activity. "Biological activity", as used herein, refers to the ability of a mutant polypeptide to inhibit hepadnavirus replication, and can be measured by the assays described below.

Other terms and definitions used herein will be understood by those of routine skill in the art. For example, by "inhibiting the replication of" is meant lowering the rate or extent of replication relative to

- 12 -

replication in the absence of a mutant polypeptide of the invention. By "into proximity with the hepadnavirus" is meant introducing into a cell, organ, or organism which is infected with a naturally-occurring hepadnavirus, or, 5 in the case of laboratory application, cotransfection or co-inoculation with a wild type hepadnavirus. By "nucleic acid" is meant deoxyribonucleic acid (DNA) or ribonucleic acid (RNA).

The methods, nucleic acids, and polypeptides of 10 the invention can be used to inhibit hepadnaviral replication in a mammal, e.g., as an effective therapy for treating individuals with a persistent HBV infection, or as a means of reducing the risk of hepatocellular carcinoma in an infected animal. Polypeptides of the 15 invention can be administered to an infected animal either directly or by gene therapy techniques. The screening methods of the invention are simple, rapid, and efficient assays designed to identify polypeptides with anti-hepadnaviral activity.

20 Other features and advantages of the invention will be apparent from the following detailed description and from the claims.

#### Brief Description of the Drawings

Fig. 1 is a schematic illustration of the 25 structural organization of "wild type" and mutant hepadnavirus constructs.

Fig. 2 is an autoradiographic image of an agarose gel, showing a Southern blot analysis of core particle DNA that was extracted from HuH-7 cells five days post 30 transfection and probed with full length <sup>32</sup>P-labeled WHV DNA.

Fig. 3 is an autoradiographic image of an agarose gel, showing a Southern blot analysis of core particle associated viral DNA that was extracted from HuH-7 cells

- 13 -

five days after transfection and probed simultaneously with the two full length  $^{32}\text{P}$ -labeled WHV and HBV DNA probes.

Fig. 4 is an autoradiographic image of an agarose gel, showing a Southern blot analysis of core particle associated viral DNA that was extracted from HepG2 cells five days after transfection and probed with full length  $^{32}\text{P}$ -labeled HBV DNA.

Fig. 5 is an autoradiographic image of a polyacrylamide gel showing a RNase protection assay.

Fig. 6 is an autoradiographic image of an agarose gel showing a Southern blot analysis of the anti-viral effect of dominant negative core mutants on "wild type" HBV replication during transient transfection in HuH-7 cells.

Fig. 7 is an autoradiographic image of a SDS-polyacrylamide gel showing a Western blot analysis of HepG2 cell lysates probed with anti-HBc antibodies.

Fig. 8 is an autoradiographic image of an agarose gel showing a Southern blot analysis of the effect of a dominant negative core mutant on replication of HBV in Hep-G2 2215 cells.

Fig. 9 is an autoradiographic image of an agarose gel showing a Southern blot analysis of cytosol-derived nucleocapsids from transfected LMC cells hybridized with a full length DHBV DNA probe.

Fig. 10 is an illustration of the nucleic acid sequence of the pCN4 plasmid insert (SEQ ID NO: 1) and the corresponding translated amino acid sequence (SEQ ID NO: 2).

Fig. 11 is an illustration of the nucleic acid sequence of the pHBV DN plasmid insert (SEQ ID NO: 3) and the corresponding translated amino acid sequence (SEQ ID NO: 4).

- 14 -

Fig. 12 is an illustration of the nucleic acid sequence of the pHBV DN AA plasmid insert (SEQ ID NO: 5) and the corresponding translated amino acid sequence (SEQ ID NO: 6).

5 Fig. 13 is an illustration of the nucleic acid sequence of the pHBV DN BB plasmid insert (SEQ ID NO: 7) and the corresponding translated amino acid sequence (SEQ ID NO: 8).

10 Fig. 14 is an illustration of the nucleic acid sequence of the pDHBV BK plasmid insert (SEQ ID NO: 9) and the corresponding translated amino acid sequence (SEQ ID NO: 10).

15 Fig. 15 is an illustration of the nucleic acid sequence of the HBV core protein (SEQ ID NO: 11) and the corresponding translated amino acid sequence (SEQ ID NO: 12).

20 Fig. 16 is an illustration of the nucleic acid sequence of the HBV core protein (SEQ ID NO: 13) and the corresponding translated amino acid sequence (SEQ ID NO: 14).

#### Detailed Description

Applicants have observed that replication of a wild type hepadnavirus is reduced when it is co-transfected with a nucleic acid construct encoding a truncated core protein, or a core-surface fusion protein. The truncated core protein, alone or in combination with the surface protein component, has a deletion of at least three amino acids from the carboxyterminal end. Viral replication is reduced by as much as 90-95% without detectable toxic effects on the host cell. Constitutively expressing a HBV mutant core-surface fusion protein as a retroviral insert substantially inhibits HBV viral DNA production in cells that previously had continuously produced all viral

- 15 -

replicative intermediates and infectious virions. An adenoviral-based plasmid that encodes the same mutant core-surface fusion protein also inhibits HBV replication following transient cotransfection in HCC cells. These  
5 dominant negative effects on viral replication are consistent over a range of hepadnavirus species.

#### Materials and Methods.

Materials and methods useful for practicing the invention are described as follows:

10        **Plasmid Constructs.** The parental plasmid pCMW82 was used to generate a series of constructs expressing WHV core proteins with an altered carboxyterminal region. Plasmid pCMW82 expresses the "wild type" WHV pregenome under the control of the cytomegalovirus immediate-early  
15 (CMV IE) promoter (Seeger et al., *J. Virol.*, **63**, 4665-4669, 1989). The pHBV plasmid carries the HBV pregenome under the control of the CMV IE promoter. These plasmids direct the synthesis of complete virions in tissue culture cells. The first nucleotide of the precore open  
20 reading frame was designated as nucleotide number 1 in the WHV genome.

The structural organization of "wild type" and mutant WHV, HBV, and DHBV core plasmid constructs are depicted in Fig. 1. The white boxes represent the open  
25 reading frame (ORF) used for constructing core mutants. Numbers at the boundaries of each ORF refer to the amino acids in the "wild type" or mutant proteins. Dotted lines represent deleted sequences. Solid and hatched boxes correspond to mutant core proteins expressed from  
30 WHV and HBV, respectively. Shaded bars refer to DHBV. The shaded hatched bars refer to the polymerase gene. Except for the "wild type" constructs pCMW82 and pCMW-DHBV, all other constructs are incapable of replication because of deletions in genes that overlap the truncated

- 16 -

portions of the core protein. The \* refers to a stop codon introduced by a frame shift mutation.

The constructs shown in Fig. 1 were produced by complete digestion with the appropriate restriction enzyme. This was followed by subsequent incubation at 30°C for 20 min. in the presence of the Klenow fragment of DNA polymerase I and deoxyribonucleotide triphosphates, which filled in the 3' recessed DNA ends. Plasmids were then ligated with T4 DNA ligase. The 3' protruding ends were filled in by incubation with the Klenow fragment of DNA polymerase I in the absence of deoxyribonucleotide triphosphates at 37°C for 15 min. This eliminated protruding ends. Deoxyribonucleotide triphosphates were then added and incubation was carried out at 30°C for 20 min. (Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989).

Constructs containing mutations in sequences encoding the core protein were obtained as follows: 1) pCN1: To make the plasmid pCN1 the WHV core gene was digested with the restriction enzyme SstI at nucleotide (nt) 310, incubated with Klenow DNA polymerase, and reclosed with T4 DNA ligase. This introduced a frame shift mutation at nt. 306 in the WHV core gene, thereby creating a stop codon at nt. 317. This mutation produces a 74 amino acid carboxyterminal truncated core protein, leaving intact the rest of the viral coding regions. 2) pCN2: To make pCN2 the WHV parental plasmid was digested with the restriction enzymes BglII (nt. 601 in the core gene) and SmaI, the latter being located in the downstream multiple cloning site of the vector. The intervening viral genes were separated by gel electrophoresis, and the DNA ends were filled in with Klenow DNA polymerase and ligated with T4 DNA ligase. This WHV core gene has 12 amino acids deleted at the

- 17 -

carboxyterminus and is fused to a three amino acid heterologous extension from the plasmid vector. 3) pCN3: To make the plasmid pCN3, the wild type plasmid pCMW82 was digested with the restriction enzymes BglII and SacII (position 2983 in the WHV X gene), the intervening viral DNA fragment was removed, and ends were filled in and ligated. The resulting plasmid construct encodes a 171 amino acid core protein fragment fused in-frame with the X protein at amino acid 31. 4) pCN4: The plasmid pCN4 was produced by a BglII-MscI (position 1826) fragment excised from pCMW82 and blunted by Klenow DNA polymerase. The plasmid was ligated to join the WHV 171 amino acid core protein as an in-frame fusion protein with amino acid 47 of the WHV small surface protein. 5) pCN5: The plasmid pCN5 was produced by removing the DNA fragment SstI (pos. 306)-BspEI (pos. 519) from pCN4, and blunting the ends with Klenow DNA polymerase and T4 DNA ligase. This introduced a WHV core in-frame deletion between amino acids 74 and 145. 6) Plasmid pCN6 expresses the first 171 amino acids of the WHV core protein fused in-frame with the HBV small surface protein at amino acid 51.

The HBV numbering system designates the unique EcoRI site as nucleotide 1. Construct pHBV DN was generated by digesting pCMW82 at nt. 601 of the core gene with BglII, and blunting the DNA end with Klenow DNA polymerase. A second cut was performed with PvuI in the ampicillin resistance gene of the carrier plasmid, and the BglII-PvuI DNA fragment was removed by fractionation on an agarose gel. The HBV MscI (pos. 299)-PvuI (in the ampicillin resistance gene of pHBV) fragment was ligated to the blunted BglII-PvuI fragment.

In order to produce an in-frame dominant negative construct of HBV that was similar to the pCN4 WHV construct, the pCN6 fragment from the SnaBI site (which

- 18 -

cuts in the CMV promoter of the carrier plasmid) to the BspEI site (pos. 519 in the WHV) was removed and substituted with the same SnaBI-BspEI (pos. 2327) fragment from pHBV. In this way, the HBV core protein

5 was fused in-frame to amino acid 144 of the WHV core protein. This fragment, derived from plasmid pCN6, was already fused at amino acid 171 to the HBV small surface protein at amino acid 51. The resulting pHBV DN therefore encodes, in the hinge between the deleted core

10 and surface proteins, five amino acids derived from the WHV core protein (GGARA). These five amino acids were not present in the subtype HBV core protein. The carboxyterminal 20 amino acid of the WHV core protein are conserved in HBV.

15 Two additional plasmids were derived from pHBV and called pHBV AA and pHBV BB. To make pHBV DN AA, pHBV was partially digested with the restriction enzyme AvaI (nt. 2431), and then partially digested with AvrII (nt. 176). The resulting DNA ends were blunted by adding Klenow DNA

20 polymerase and nucleotide triphosphates. The DNA ends were ligated with T4 ligase. The resulting plasmid pHBV DN AA encodes the HBV core protein fused in frame at amino acid 179 with the surface protein (encoded by the "S gene") at amino acid 8. The plasmid pHBV BB was made

25 by performing two sequential partial digestions with the enzymes BglII and BamHI. The DNA ends were ligated with T4 ligase. The pHBV BB plasmid expresses the HBV core protein fused in frame at amino acid 175 with the surface protein at amino acid 112. The correct design of the

30 constructs was confirmed by restriction digest mapping and DNA sequence analysis. Plasmid DNAs were purified by the alkali lysis procedure followed by sedimentation through a cesium chloride-ethidium bromide density gradient. As a result of these viral gene manipulations

35 the above plasmid constructs produce replication

- 19 -

defective WHV genomes. Plasmid pCN1 expresses a truncated core protein that is unable to assemble into functional nucleocapsids. All other constructs contain inactivating deletions in the polymerase gene.

5 Another plasmid, designated pRHBBE, was constructed using the polylinker of the plasmid pBS SK(+) (Stratagene), which allows for viral gene transcription from the T7 promoter to make a HBV-specific 276 nt antisense RNA. This species, encoded by a BamHI  
10 (pos. 2906) to EcoRI (pos. 1) fragment, was used in RNase protection experiments. The <sup>32</sup>P labeled riboprobe annealed specifically to the "wild type" pregenomic HBV DNA without recognizing the pHBV DN mRNA.

Constructs expressing DHBV dominant negative  
15 proteins were derived from the plasmid pCMV DHBV (Wu et al., *J. Virol.*, **65**, 2155-2163, 1991), which expresses the DHBV pregenome under the control of the CMV promoter. Construct pSK contains a deletion between the SphI site (position 2843 in the core gene; this numbering system is  
20 arbitrarily initiated with the nucleotide GAATTC of the unique EcoRI site) and the KpnI site (position 1290, in the S gene). The intervening fragment was separated by agarose gel electrophoresis. The ends of the larger DNA fragment were blunted by Klenow DNA polymerase and  
25 religated. This construct expresses, under the control of the CMV promoter, a protein composed of the first 66 amino acids of the DHBV core protein fused in frame to amino acid five of the DHBV surface protein. Construct pBK contains a deletion between the BglII site (position  
30 391 in the core gene) and the KpnI site (position 1290 in the S gene). The intervening fragment was separated by agarose gel electrophoresis and the ends of the larger DNA fragment were filled in and blunted by the Klenow DNA polymerase. The ends were then religated. The resulting  
35 construct expresses, under the control of the CMV

- 20 -

promoter, a protein composed of the first 257 amino acids of the DHBV core protein (five amino acids are missing from the carboxyterminus), fused in frame to the fifth amino acid of the DHBV surface protein.

5       To make the construct pK, the pCMV DHBV was linearized by cutting at the KpnI site (position 1290 of the S gene). The DNA ends were blunted with the Klenow DNA polymerase reaction and the fragment was religated. The resulting construct has a frame-shift mutation so  
10 that the DHBV polymerase pK gene and the pre-S and S genes have a termination site a few nucleotides downstream from the KpnI site. The construct pK thus expresses, under the control of the CMV promoter, the full length core protein, but none of the envelope  
15 proteins apart from a truncated pre-S protein. A frameshift mutation that occurs in the polymerase gene renders the other constructs carrying the deletions described above replication defective. Construct pNX contains a deletion between the NsiI site (position 2845  
20 in the core gene) and the XhoI site (position 1212 in the pre-S gene). The intervening fragment was separated by agarose gel electrophoresis. The ends of the larger DNA fragment were blunted and filled in with Klenow DNA polymerase, followed by religation of the fragment to  
25 itself. This construct expresses, under the control of the CMV promoter, the first 68 amino acids of the DHBV core protein fused in frame to amino acid 437 of the carboxyterminus of the polymerase gene.

      Retroviral constructs: The HBV core-surface  
30 fusion gene encoded by pHBV DN was PCR amplified with oligonucleotides containing at their 5' ends a SalI restriction enzyme recognition site. The antisense primer contained a recognizable Flag epitope (Kodak). The PCR product was gel purified, digested with SalI, and  
35 cloned in the retroviral pBabe Puro vector (Morgenstern

- 21 -

et al., *Nucl. Acids Res.*, 18:3587-96, 1990) at its Sall site. The design of the resulting pBP HBV DN vector was confirmed by sequence analysis.

- Transfections into Hepatoma Cell Lines: Human
- 5 hepatoma cells HuH-7 and HepG2 support a complete viral replicative cycle following transfection with a plasmid construct expressing the pregenomic viral RNA (Mason et al., *Hepatology*, 2, 635-45, 1989). Cells were maintained and passaged as previously described (Wu et al., *J.*
- 10 *Virology*, 65, 2155-2163, 1991). Cells were transiently co-transfected with plasmids expressing the mutated WHV or HBV core genes (described above), together with an equal amount of a "wild type" WHV or HBV producing plasmid. Co-transfections were performed by the calcium phosphate
- 15 technique (CaPO<sub>4</sub> transfection Kit, 5'-3', Boulder, Colorado). Briefly,  $1.2 \times 10^7$  cells in 100 mm plates were grown for 24 hours. The medium was changed 2-4 hours before transfecting with 10  $\mu$ g of "wild type" virus. This produces the plasmid along with 10  $\mu$ g of
- 20 each mutant construct. The precipitate was left on the cells for 6-8 hours, and then the medium was replaced. The cells were harvested two days after transfection when performing RNA experiments, and five days after transfection when performing DNA experiments.
- 25 The cell line LMH, derived from a chicken hepatocellular carcinoma, was used for transfection of the DHBV derived plasmids. This cell line supports higher levels of DHBV replication than do cell lines of human origin. Another cell line, derived from LMH and
- 30 named D2, was created by stably transfecting a head-to-tail DHBV dimer that produces infectious DHBV particles. These cells were grown in DMEM and 10% FCS and transfected with the various dominant negative core mutant constructs as described above.

- 22 -

- Infection of the HepG2 2215 Cell Line:** Infection of the HepG2 2215 cell line by recombinant retroviruses was accomplished following a standard protocol for producing retroviral stocks and for infecting tissue
- 5 culture cells (Miller et al., *Biotechniques*, 7, 980-990, 1989; Miller, et al., *Methods in Enzymology*, 217, 581, 1993). After infection, the cells were selected with 2  $\mu$ g/ml puromycin (Sigma). Resistant clones were pooled and further expanded.
- 10 **Analysis of viral DNA replication.** WHV and HBV DNA replication were assayed by Southern blot analysis of DNA that had been extracted from intracellular core particles. The procedure for isolation of core particles was previously described in detail (Pugh et al., *J.*
- 15 *Virology*, 62, 3513-3516, 1988). DNA fractionation on agarose gels was performed under alkali conditions and the DNA was transferred onto a nylon membrane for Southern blot analysis (Hybond N, Amersham International, Little Chalfont, UK). Prehybridization and hybridization
- 20 reactions were carried out at 65°C in 6X SSC solution (1X SSC is 150 mM NaCl, 15 mM Na<sub>3</sub>Citrate), 5X Denhardt's solution (100X is 2% w/v BSA, 2% w/v Ficoll, 2% w/v polyvinyl pyrrolidone), and 0.5% SDS. WHV and HBV DNAs were detected by hybridization with randomly primed <sup>32</sup>P-
- 25 labeled full length WHV or HBV DNA (Multiprime DNA Labelling System, Amersham). The membranes were washed twice for 15 min. each at 65°C in 1X SSC, 0.1% SDS, and were then washed once more at 65°C in 0.1X SSC, 0.1% SDS. The nylon membranes were then autoradiographed at -70°C,
- 30 using intensifying screens and Kodak films. Signal intensities on the nylon sheets were quantitated by a computer assisted scanning system (Ambis Quantprobe System version 3.0).

- Extraction and analysis of viral RNA.** Total RNA
- 35 was extracted from a 100 mm dish by lysis of cells in 1

- 23 -

ml of solution D (4 M guanidinium thiocyanate, 25 mM NaCitrate, pH 7.0, 0.5% sarcosyl, 0.1 M 2-mercaptoethanol) as described (Chomczynski et al., Anal. Biochem., 168, 156-159, 1987). Encapsidated viral RNA was  
5 extracted from core particles by lysis in 200 ml of solution D and the volumes were adjusted accordingly as described (Roychoury et al., *supra*). Finally, to exclude contamination by plasmid DNA or reversed transcribed HBV DNA, the encapsidated viral RNA was digested with 16 U  
10 RNase-free DNase RQ1 DNase (Promega, Madison, WI) at 37°C for 15 min., followed by phenol-chloroform extraction and ethanol precipitation, before undergoing the RNase protection assay.

RNase protection analysis of total and  
15 encapsidated viral RNA was performed with a commercially available kit according to the manufacturer's instructions (RPA II-Ribonuclease protection kit, Ambion Inc. Austin, TX). The RNA probe was derived from the plasmid pRHBBE, a derivative of the pBluescript SK(+),  
20 which includes the 280 bp HBV fragment BamHI (pos. nt 2901)-EcoRI (pos. nt 1), oriented to produce an antisense RNA molecule when transcription was initiated with the bacteriophage T7 RNA polymerase. The RNA probe contained approximately 50 nt of plasmid sequences that were not  
25 protected by the HBV specific RNA. Labeled RNA was synthesized as follows: 0.5 µg of pRHBBE was cut by BamHI and then transcribed by T7 RNA polymerase (Promega, Madison, WI) in the presence of  $\alpha$ -<sup>32</sup>P UTP (100 µCi at 400 Ci/mM, New England Nuclear, Boston, MA). The antisense  
30 RNA probe recognized pregenomic RNA and the 2.4 pre-S1 mRNA derived from "wild type" HBV, but did not recognize transcripts derived from pHBV DN. Hybridization, after denaturation at 95°C for 3 min., was performed in 20 µl on 2 µg of total RNA or encapsidated pregenomic RNA  
35 derived from half of a 100 mm plate at 42°C overnight in

- 24 -

a solution of 80% formamide, 100 mM NaCitrate pH 6.4, 300 mM NaAcetate pH 6.4, and 1mM EDTA. RNase digestion was carried out with RNase A (0.5U) and RNase T1 (20 U) at 37°C for 30 min. Fragments protected by RNase digestion were separated on a denaturing 6% polyacrylamide gel (Sequagel 6%, National Diagnostics, Atlanta, GA).

**Viral nucleocapsid isolation and Western blots.**

HepG2 cells that had been transfected with pHBV alone, pHBV DN together, or pHBV DN alone were lysed in 500 ml TNE, 1% NP 40. The debris was pelleted by centrifugation at 10,000 rpm in an Eppendorf bench top centrifuge. A 200 µl aliquot of the clarified cell lysate was ultracentrifuged at 500,000 xg for 1 hour at 4°C through 2 ml of a 20% w/v sucrose/TNE cushion using a TLA 100 rotor (Beckman Instruments, Palo Alto, CA). Under these conditions viral core particles were pelleted, whereas free core protein and soluble hepatitis Be antigen (HBcAg) remained in the supernatant (Zhou et al., *supra*). The pellet was resuspended in 100 µl of Laemmli sample buffer and boiled for 3 min. One-half of the sample was run over a 12.5% SDS-PAGE gel (Acrygel National Diagnostics, Atlanta, GA). Western blotting was performed on an Immobilon-P membrane (Millipore Co., Bedford, MA) (Harlow et al., Antibodies: a laboratory manual, Cold Spring Harbor Laboratories, CSH, NY 1988). After transfer the membrane was blocked for one hour in a solution of 5% non-fat dry milk and 0.5% Tween-20 in phosphate buffered saline (PBS). HBcAg antigenicity was detected by incubation of the membrane with polyclonal antibodies prepared in rabbits against recombinant HBcAg (Dake Co., Carpinteria, CA) at a 1:250 dilution in the above solution for one hour at 20°C. The filter was washed at 20°C in PBS, 0.5% Tween-20 with three successive changes of solution. Bound antibody was detected using the chemiluminescence method (ECL,

- 25 -

Amersham International, Little Chalfont, UK) using peroxidase labeled goat anti-rabbit antibodies. The filter was exposed to Kodak films for 5-20 seconds.

#### Experimental Results

5           Inhibition of WHV DNA synthesis. WHV core mutant plasmids were tested for the ability to inhibit wild type WHV DNA replication in HuH-7 cells. Fig. 2 shows a Southern blot analysis of core particle DNA that was  
10 and probed with full length <sup>32</sup>P labeled WHV DNA. Lane M contains <sup>32</sup>P 5' end labeled lambda HindIII molecular weight markers. The HuH-7 cells were transfected with: lane 1, pCMW82; lane 2, pCMW82 and pCN4; lane 3, pCMW82 and pCN1; lane 4, pCMW82 and pCN2; lane 5, pCMW82 and  
15 pCN3; and lane 6, pCMW82 and pCN5. Each lane was loaded with one-half of the core associated viral DNA, which had been extracted from a 100 mm tissue culture dish of HuH-7 cells.

          All mutant WHV core constructs suppressed "wild  
20 type" WHV DNA synthesis, albeit with different efficiencies. The extent of inhibition varied among the different constructs, depending in part on the molecular structure of the mutant core protein. In order to exclude experimental variability, all transfections were  
25 repeated several times with comparable results. The data represent an average of at least three independent experiments. Cotransfection of "wild type" pCMW82 with the mutant core constructs pCN1, pCN2, and pCN3 produced a modest inhibition of "wild type" viral DNA replication  
30 (36%, 48%, and 12%, respectively). In contrast, pCN4 and pCN5 substantially inhibited WHV DNA synthesis in HuH-7 cells by 90% and 85%, respectively (Fig. 2).

          To test whether the pCN4 construct inhibits HBV replication, cotransfection experiments were performed

- 26 -

with "wild type" pHBV. There was no reduction of HBV synthesis by the WHV based construct pCN4. Fig. 3 shows a Southern blot analysis of core particle associated viral DNA extracted from HuH-7 cells five days after

5 transfection. The blots were probed simultaneously with full length  $^{32}\text{P}$  labeled WHV and HBV DNA probes. Lane M contains  $^{32}\text{P}$  5' end labeled lambda HindIII molecular weight markers. Core particle associated viral DNA was extracted from cells transfected with: lane 1, pCMW82;

10 lane 2, pCMW82 and pCN4; lane 3, pCMW82 and pCN6; lane 4, pHBV; lane 5, pHBV and CN4; and lane 6, pHBV and pCN6. Each lane was loaded with one-half of the core particle associated DNA that had been extracted from a 100 mm tissue culture dish of HuH-7 cells.

15 In order to determine the general region of the fusion protein that was responsible for inhibiting viral replication, a chimeric construct expressing WHV core-HBV small surface fusion protein was produced. This plasmid, designated pCN6, reduced "wild type" WHV replication by

20 85%, an inhibitory effect comparable to the original parental construct pCN4. Like pCN4, pCN6 does not inhibit HBV replication (Fig 3, lane 6). It was concluded that the WHV core-small surface fusion protein encoded by pCN4 exerts a species-specific inhibitory

25 effect.

To determine the amount of pCN4 required to interfere effectively with WHV replication, HuH-7 cells were co-transfected at various ratios of CMW82 to pCN4 using 10  $\mu\text{g}$  of pCMW82. The total amount of transfected

30 DNA was kept constant (20  $\mu\text{g}$ ) by adding unrelated sonicated salmon sperm DNA. The results of these experiments indicate that when pCN4 was diluted by 10 and 50 fold, there was still a 66% and 20% inhibition of "wild type" WHV replication, respectively. Interference

- 27 -

with viral replication occurs even in the presence of an excess of "wild type" core protein.

Dominant negative core mutant polypeptides are not toxic to HCC cells. To insure that the mutant plasmids were neither affecting the efficiency of DNA uptake by HuH-7 cells during transfection, nor inducing a cytopathic effect, each 100 mm plate had a 10 mm cover slip containing cells grown under the same conditions. The cells were investigated by immunocytochemistry utilizing the protocol of Jilbert et al. (*J. Virol.*, 66, 1377-1378, 1992). Core protein expression was detected with polyclonal antibodies prepared against either WHV or HBV recombinant core proteins. Approximately one percent of the HuH-7 cells were transfected with the "wild type" WHV plasmid, as demonstrated by diffuse cytoplasmic staining for WHcAg in cells harvested five days post transfection. After transfection of cells with pCN4 alone, a punctate distribution of WHcAg in the perinuclear region was observed. The same staining pattern was obtained when the dominant negative core mutant constructs were co-transfected with "wild type" pCMW82. The total number of HBcAg positive cells did not vary under these conditions. The mutant core expressing plasmids did not inhibit "wild type" viral DNA uptake during the transfection process and were not toxic to HuH-7 cells.

It was also necessary to exclude the possibility that the inhibitory effect exerted by pCN4 on WHV replication was the result of decreased transcription of "wild type" WHV pregenomic RNA. For these studies, Poly(A)<sup>+</sup>RNA was extracted from HuH-7 cells that had been transfected with the plasmids pCMW82 alone, pCMW82 and pCN4 together, or pCN4 alone. The RNA was probed with a BglII-BstXI WHV DNA fragment that specifically recognized the pregenomic RNA but not the pCN4 transcripts. The

- 28 -

results demonstrated no change in the level of "wild type" WHV pregenomic RNA transcription from pCMW82 in the presence of pCN4.

Inhibition of HBV replication. Based on the previous results, it was of interest to determine whether a similar mutant core polypeptide would inhibit HBV replication in HCC cells. The construct pHBV DN was designed to be the molecular HBV-derived equivalent of pCN4. Plasmid pHBV DN was co-transfected with "wild type" pHBV into HuH-7 and HepG2 cells. It inhibited "wild type" HBV DNA replication by 90% (Fig. 4).

Fig. 4 shows a Southern blot analysis of core particle associated viral DNA extracted from HepG2 cells five days after transfection. The blot was probed with full length <sup>32</sup>P labeled HBV DNA. Lane M contains <sup>32</sup>P 5' end labeled lambda HindIII molecular weight markers. Lane 1 contains 3.2 kb linear HBV DNA (10 pg). The remaining lanes show core particle associated viral DNA extracted from cells transfected with pHBV (lane 2); or pHBV and pHBV DN (lane 3).

The constructs pHBV DN AA and pHBV DN BB were assayed in the same manner, for the purpose of mapping which regions of the core protein and of the surface protein were necessary for inhibition. The construct pHBV DN AA was at least as potent an inhibitor as pHBV DN, whereas pHBV DN BB was less inhibitory than pHBV DN. This is shown in Fig. 5, which is a Southern blot analysis illustrating the antiviral effects of the pHBV DN AA and pHBV DN BB dominant negative core mutants on "wild type" HBV replication during transient transfection experiments in HuH-7 cells. The pCMV-HBV lane shows the level of "wild type" HBV replication in HUH-7 cells. The dominant negative mutant pHBV-DN reduced wild type replication by 95%. When this construct was placed in a vector containing the adenovirus sequences necessary for

- 29 -

producing a recombinant adenovirus vector (Ad HBV DN), there was an 80% decrease in HBV replication. When the HBV DN construct was placed in a retroviral vector (e.g., pBP HBV DN), there was a 90-95% reduction in HBV replication.

Experiments were then performed to assess the presence and amount of pregenomic RNA within nucleocapsids, and to compare these results to the level of viral RNA present in the cytosolic fraction by means of a sensitive RNase protection assay (Fig. 6). RNA was extracted from HepG2 transfected cells and probed with a <sup>32</sup>P labeled 322 nt RNA probe containing the BamHI (pos. 2906)-EcoRI (pos. 1) fragment (lane P), or electrophoresed on a 6% denaturing PAGE gel after RNase A and T1 digestion. Lane 1 contains 2 µg of total RNA from HepG2 cells transfected with pHBV; lane 2 contains 2 µg of total RNA from HepG2 cells transfected with pHBV and pHBV DN; lane 3 contains 2 µg of total RNA from HepG2 cells transfected with pHBV DN alone (the BamHI-EcoRI fragment is missing in this construct). The remaining lanes show RNA that was extracted from HepG2-derived core particles and then probed as in lanes 1-3. Each lane was loaded with half of the core associated RNA extracted from a 10 cm dish. Lane 4 contains core particle associated RNA from cells transfected with pHBV. Lane 5 contains core particle associated RNA from cells transfected with pHBV and pHBV DN. Lane 6 contains core particle associated RNA from cells transfected with pHBV DN alone. There was a 90% reduction in encapsidation of "wild type" pregenomic RNA when pHBV DN was co-transfected with the wild type HBV DNA expressing plasmid, whereas no significant reduction in viral RNA was observed in experiments performed with total cellular RNA. The riboprobe used in this experiment protects

- 30 -

pregenomic and pre-S1 transcripts, both of which were absent in the pHBV DN transfected cells.

"Wild type" pregenomic viral RNA is incapable of being encapsidated in the presence of mutant core protein because of inefficient core particle assembly. Cell lysates derived from HepG2 cells previously transfected with pHBV alone, pHBV and pHBV DN together, and pHBV DN alone were sedimented on a 20% w/v sucrose cushion for one hour at 500,000 g. Under these experimental conditions non-particulate core protein and HBeAg were left in solution (Zhou et al., *supra*). The pellet was analyzed for core protein by 12.5% SDS-PAGE electrophoresis, and analyzed on a Western blot using polyclonal anti-HBc antibodies as probes (Fig. 7). The viral core particles were derived from: lane 1, cells transfected with pHBV; lane 2, cells transfected with pHBV and pHBV DN; lane 3, cells transfected with pHBV DN alone; lane 4, HepG2 2215 cells (positive control). Lane 5 contains 100 µg of cell lysate in RIPA buffer not subjected to ultracentrifugation and extracted from HepG2 2215 cells to show enrichment of core particles by the pelleting technique (positive control). The protein in lane 6 was derived from the pellets of untransfected HepG2 cells (negative control). A protein band of 21.5 kd, corresponding to the intact "wild type" HBV core protein, was detected only in the pellet derived from HepG2 cells transfected with pHBV. In the pellet of cells transfected with the pHBV DN plasmid, an immunoreactive core protein band of 11.5 kd was detected. This protein was substantially smaller than the predicted size of the full length core-surface fusion protein derived from the pHBV DN (about 38 kd).

To determine whether the HBV core dominant negative mutant HBV DN can make hepatoma cell lines resistant to HBV replication, the HBV DN coding sequence

- 31 -

was cloned into the retroviral vector pBabe Puro (pBP), which contains a puromycin selectable marker. The resulting vector is named pBPHBV DN. Recombinant retroviral stocks were obtained after transfecting pBPHBV DN into the packaging cell line PA317. The stocks were then used to infect HepG2 and HepG2 2215 cell lines. The HepG2 2215 cells constitutively produce wild type HBV virions due to the stable integration of a head to tail dimer of HBV. Pools of stably transduced clones were grown in the presence of puromycin. HBV DNA was purified from the core particles and analyzed by Southern blot. HepG2 2215 transduced by the pBP HBV DN vector showed a 90% reduction in HBV replication when compared to HepG2 2215 cells transduced by the pBP vector (Fig. 8). This result demonstrates a striking reduction of HBV replicative intermediates in core particles, even in a cell line that constitutively expresses all the viral gene products and replicative forms of the virus.

The Flag tagged dominant negative form of the HBV DN sequence was also cloned into the adenoviral vector pAdBgIII to generate the vector pAdHBV DN. This vector contains a multiple cloning site flanked by the CMV EI promoter and by adenovirus 5 sequences. The adenovirus 5 sequences allow homologous recombination and reconstitution of a recombinant replication incompetent adenovirus after cotransfection in 293 cells (Graham et al., the Human press, Vol 7, 109-128, 1991). The plasmid pAdHBV DN was then introduced, along with pHBV, into HCC cells by transient transfection, inhibiting HBV replication by 80% (Fig. 5). Adenoviral vectors such as pAd HBV DN can be used to generate a replication incompetent adenovirus by homologous recombination, and can express the HBV DN polypeptide in the liver.

**Inhibition of DHBV replication** Substantial suppression of DHBV replication was obtained by co-

- 32 -

transfecting pCMV DHBV with the plasmid pBK. The pBK plasmid encodes a DHBV core protein which lacks the last carboxyterminal five amino acids, fused to a surface protein which lacks the aminoterminal first four amino acids. When shorter core fragments were fused in frame to a surface protein lacking the aminoterminal first four amino acids (plasmid pSK), to the Pol gene product (pNX), or to the pre-S gene product (pSK), there was little or no effect on DHBV replication. This result indicated that both the core protein and the surface protein extension were important for exerting an inhibitory effect on "wild type" DHBV replication, presumably by disrupting nucleocapsid assembly. The core portion of the chimeric mutant polypeptide interacts with the wild type core protein, preventing formation of intact nucleocapsids and thus encapsidation of the DHBV pregenome. A construct that expressed only the DHBV core protein (pK) was incapable of inhibiting DHBV replication, while a plasmid that expressed the same core portion as the pBK plasmid but fused to the polymerase gene was incapable of inhibiting "wild type" DHBV replication. Fig. 9 is a Southern Blot analysis of cytosolic derived nucleocapsid DNA from transfected LMC cells, hybridized to a full length DHBV DNA probe. LMC cells were transfected with 10  $\mu$ g of pCMV DHBV together with 10  $\mu$ g of mutant plasmids pSK (lane 2), pBK (lane 3), pSK (lane 4, the same as lane 2), pK (lane 5), or pNX (lane 6). The last lane contains the cytosolic derived nucleocapsid DNA from a LMC cell line stably transfected with a head-to-tail DHBV dimer as a positive control. Replication of "wild type" DHBV was inhibited by the dominant negative core mutant construct BK.

- 33 -

Therapeutic Use

The mutant polypeptides of the invention can be provided exogenously to a target cell of an animal suspected of harboring a hepadnavirus infection by any appropriate method, for example by oral, parenteral, transdermal, or transmucosal administration. The mutant polypeptide can be administered in a sustained release formulation using a biodegradable biocompatible polymer, or by on-site delivery using micelles, gels or liposomes. Therapeutic doses can be, but are not necessarily, in the range of 0.01 - 100.0 mg/kg body weight, or a range that is clinically determined to be appropriate by those skilled in the art.

The polypeptides useful in a method of the invention, or as candidate agents in a method of the invention, can be purified using conventional methods of protein isolation known to one skilled in the art. These methods include, but are not limited to, precipitation, chromatography, immunoabsorption, or affinity techniques (see, e.g., Scopes, R. Protein Purification: Principles and Practice, 1982 Springer Verlag, NY). The polypeptide can be purified from starting material that is derived from a genetically engineered cell line. One useful method of purification involves expressing the polypeptide as a fusion protein encoded by a glutathione-S-transferase vector, purifying the resulting fusion protein by GST-GSH affinity chromatography, and removing the GST portion of the fusion polypeptide by thrombin cleavage. Alternatively, a synthetic mutant polypeptide can be prepared by automated peptide synthesis (see, e.g., Ausubel et al., eds. Current Protocols in Molecular Biology, John Wiley & Sons, publ. NY. 1987, 1989; Sambrook et al. (1989), Molecular Cloning: A Laboratory Manual (2d ed.), CSH Press).

- 34 -

Therapeutic administration of a mutant polypeptide can also be accomplished using gene therapy techniques. A nucleic acid that included a promoter operatively linked to a sequence encoding a polypeptide of the invention is used to generate high-level expression of the polypeptide in cells transfected with the nucleic acid. Gene transfer can be performed *ex vivo* or *in vivo*. To administer the nucleic acid *ex vivo*, cells can be removed from the body of the patient, transfected with the nucleic acid encoding the mutant polypeptide, and returned to the patient's body. Alternatively the nucleic acid can be administered *in vivo*, by transfecting the nucleic acid into target cells (e.g., hepatocytes) so that the mutant polypeptide is expressed *in situ*.

The nucleic acid molecule is contained within a non-replicating linear or circular DNA or RNA molecule, or within an autonomously replicating plasmid or viral vector, or may be integrated into the host genome. Any vector that can transfect a cell can be used in the methods of the invention. Preferred vectors are viral vectors, including those derived from replication-defective hepatitis virus (e.g., HBV and HCV), retrovirus (see, e.g., WO89/07136; Rosenberg et al., N. Eng. J. Med. 323(9):570-578, 1990; Miller et al., 1993 *supra*), adenovirus (see, e.g., Morsey et al., J. Cell. Biochem., Supp. 17E, 1993; Graham et al., in Murray, ed., Methods in Molecular Biology: Gene Transfer and Expression Protocols. Vol. 7, Clifton, NJ: the Human Press 1991: 109-128), adeno-associated virus (Kotin et al., Proc. Natl. Acad. Sci. USA 87:2211-2215, 1990), replication defective herpes simplex virus (HSV; Lu et al., Abstract, page 66, Abstracts of the Meeting on Gene Therapy, Sept. 22-26, 1992, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York), and any modified versions of these vectors. Other preferred viral vectors include those

- 35 -

modified to target a specific cell type (see, e.g., Kan et al. WO 93/25234; Kasahara et al. Science, 266:1373-76, 1994; Dornburg et al. WO 94/12626; Russell et al. WO 94/06920). Methods for constructing expression vectors  
5 are well known in the art (see, e.g., Molecular Cloning: A Laboratory Manual, Sambrook et al., eds., Cold Spring Harbor Laboratory, 2nd Edition, Cold Spring Harbor, New York, 1989).

Appropriate regulatory sequences can be inserted  
10 into the vectors of the invention using methods known to those skilled in the art, e.g., by homologous recombination (Graham et al., J. Gen. Virol. 36:59-72, 1977), or by other appropriate methods (Sambrook et al., eds., supra). Promoters are inserted into the vectors so  
15 that they are operatively linked 5' to the nucleic acid sequence encoding the mutant polypeptide. Any promoter that is able to initiate transcription in a target cell can be used in the invention. For example, non-tissue specific promoters, such as the cytomegalovirus  
20 (DeBernardi et al., Proc. Natl. Acad. Sci. USA 88:9257-9261, 1991, and references therein), mouse metallothionine I gene (Hammer, et al., J. Mol. Appl. Gen. 1:273-288, 1982), HSV thymidine kinase (McKnight, Cell, 31:355-365, 1982), and SV40 early (Benoist et al.,  
25 Nature, 290:304-310, 1981) promoters may be used. Preferred promoters for use in the invention are hepatocyte-specific promoters, the use of which ensures that the mutant polypeptides are expressed primarily in hepatocytes. Preferred hepatocyte-specific promoters  
30 include, but are not limited to the albumin, alpha-fetoprotein, alpha-1-antitrypsin, retinol-binding protein, and asialoglycoprotein receptor promoters. Additional viral promoters and enhancers, such as those from herpes simplex virus (types I and II), hepatitis  
35 virus (Types A, B, and C), and Rous sarcoma virus (RSV;

- 36 -

Fang et al., Hepatology 10:781-787, 1989), can also be used in the invention.

The mutant polypeptides of the invention, and the recombinant vectors containing nucleic acid sequences encoding them, may be used in therapeutic compositions for preventing or treating HBV infection. The therapeutic compositions of the invention may be used alone or in admixture, or in chemical combination, with one or more materials, including other mutant polypeptides or recombinant vectors, materials that increase the biological stability of the oligonucleotides or the recombinant vectors, or materials that increase the ability of the therapeutic compositions to penetrate hepatocytes selectively. The therapeutic compositions of the invention can be administered in pharmaceutically acceptable carriers (e.g., physiological saline), which are selected on the basis of the mode and route of administration, and standard pharmaceutical practice. Suitable pharmaceutical carriers, as well as pharmaceutical necessities for use in pharmaceutical formulations, are described in *Remington's Pharmaceutical Sciences*, a standard reference text in this field.

The therapeutic compositions of the invention can be administered in dosages determined to be appropriate by one skilled in the art. An appropriate dosage is one which effects a reduction in a disease caused by HBV infection. It is expected that the dosages will vary, depending upon the pharmacokinetic and pharmacodynamic characteristics of the particular agent, and its mode and route of administration, as well as the age, weight, and health (including renal and hepatic function) of the recipient; the nature and extent of the disease; the frequency and duration of the treatment; the type of, if any, concurrent therapy; and the desired effect. It is expected that a useful dosage contains between about 0.1

- 37 -

to 100 mg of active ingredient per kilogram of body weight. Ordinarily a dosage of 0.5 to 50 mg, and preferably, 1 to 10 mg of active ingredient per kilogram of body weight per day given in divided doses, or in  
5 sustained release form, is appropriate.

The therapeutic compositions of the invention may be administered to a patient by any appropriate mode, e.g., parenterally, as determined by one skilled in the art. Alternatively, it may be necessary to administer  
10 the treatment surgically to the target tissue. The treatments of the invention may be repeated as needed, as determined by one skilled in the art.

The invention also includes any other methods which accomplish *in vivo* transfer of nucleic acids into  
15 target cells. For example, the nucleic acids may be packaged into liposomes, non-viral nucleic acid-based vectors, erythrocyte ghosts, or microspheres (microparticles; see, e.g., U.S. Patent No. 4,789,734; U.S. Patent No. 4,925,673; U.S. Patent No. 3,625,214;  
20 Gregoriadis, Drug Carriers in Biology and Medicine, pp. 287-341 (Academic Press, 1979)). Further, delivery of mutant polypeptides be accomplished by direct injection of their nucleic acid coding sequences into target tissues, for example, in a calcium phosphate precipitate  
25 or coupled with lipids, or as "naked DNA".

Mutant core polypeptides and core-surface fusion proteins of the invention can be tested for their ability to inhibit hepadnavirus replication in an animal model. For example, candidate polypeptides can be injected into  
30 an animal that is infected with a hepadnavirus, e.g., a woodchuck, duck, or ground squirrel harboring its respective hepatitis B virus variants (see, e.g., Mason et al., *J. Virol.* 36:829, 1980; Schodel et al., in Molecular Biology of hepatitis B virus, CRC press, Boca  
35 Raton, p. 53-80, 1991; Summers et al., *Proc. Natl. Acad.*

- 38 -

Sci. USA, 75:4533-4537, 1978). Candidate polypeptides can also be analyzed in transgenic animal strains developed for the purpose of studying hepadnaviral gene expression (see, e.g., Babinet et al., Science, 230:1160-5 63, 1985; Burk et al., J. Virol. 62:649-54, 1988; Chisari et al., Science 230:1157-60, 1985; Chisari, in Current Topics in Microbiology and Immunology, p. 85-101, 1991). Candidate polypeptides of the invention can also be tested in animals that are naturally infected with HBV, 10 e.g., in chimpanzees, by administering the polypeptides, or the nucleic acids encoding them, to the animal by one of the methods discussed above, or by other standard methods.

#### Other Embodiments

15 From the above description, one skilled in the art can easily ascertain the essential characteristics of the present invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various 20 usages and conditions.

All publications cited herein are fully incorporated by reference in their entirety.

Other embodiments are within the claims set forth below.

- 39 -

TABLE 1  
CONSERVATIVE AMINO ACID REPLACEMENTS

For Amino Acid	Code	Replace With
Alanine	A	D-Ala, Gly, Aib, $\beta$ -Ala, Acp, L-Cys, D-Cys
Arginine	R	D-Arg, Lys, D-Lys, homo-Arg, D-homo-Arg, Met, Ile, D-Met, D-Ile, Orn, D-Orn
Asparagine	N	D-Asn, Asp, D-Asp, Glu, D-Glu, Gln, D-Gln
Aspartic Acid	D	D-Asp, D-Asn, Asn, Glu, D-Glu, Gln, D-Gln
Cysteine	C	D-Cys, S-Me-Cys, Met, D-Met, Thr, D-Thr
Glutamine	Q	D-Gln, Asn, D-Asn, Glu, D-Glu, Asp, D-Asp
Glutamic Acid	E	D-Glu, D-Asp, Asp, Asn, D-Asn, Gln, D-Gln
Glycine	G	Ala, D-Ala, Pro, D-Pro, Aib, $\beta$ -Ala, Acp
Isoleucine	I	D-Ile, Val, D-Val, Adaa, AdaG, Leu, D-Leu, Met, D-Met
Leucine	L	D-Leu, Val, D-Val, Adaa, AdaG, Leu, D-Leu, Met, D-Met
Lysine	K	D-Lys, Arg, D-Arg, homo-Arg, D-homo-Arg, Met, D-Met, Ile, D-Ile, Orn, D-Orn
Methionine	M	D-Met, S-Me-Cys, Ile, D-Ile, Leu, D-Leu, Val, D-Val
Phenylalanine	F	D-Phe, Tyr, D-Thr, L-Dopa, His, D-His, Trp, D-Trp, Trans-3,4, or 5-phenylproline, Adaa, AdaG, cis-3,4, or 5-phenylproline, Bpa, D-Bpa
Proline	P	D-Pro, L-I-thioazolidine-4-carboxylic acid, D-or L-1-oxazolidine-4-carboxylic acid (Kauer, U.S. Patent (4,511,390))
Serine	S	D-Ser, Thr, D-Thr, allo-Thr, Met, D-Met, Met(O), D-Met(O), L-Cys, D-Cys
Threonine	T	D-Thr, Ser, D-Ser, allo-Thr, Met, D-Met, Met(O), D-Met(O), Val, D-Val
Tyrosine	Y	D-Tyr, Phe, D-Phe, L-Dopa, His, D-His
Valine	V	D-Val, Leu, D-Leu, Ile, D-Ile, Met, D-Met, Adaa, AdaG

- 40 -

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: The General Hospital Corporation
- (ii) TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
- (iii) NUMBER OF SEQUENCES: 14
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Fish & Richardson P.C.
  - (B) STREET: 225 Franklin Street
  - (C) CITY: Boston
  - (D) STATE: MA
  - (E) COUNTRY: USA
  - (F) ZIP: 02110-2804
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: 60/017,814
  - (B) FILING DATE: 20-JUN-1995
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Clark, Paul T.
  - (B) REGISTRATION NUMBER: 30,162
  - (C) REFERENCE/DOCKET NUMBER: 00786/282001
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 617/542-5070
  - (B) TELEFAX: 617/542-8906
  - (C) TELEX: 200154

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1041 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG	GAC	ATA	GAT	CCC	TAT	AAA	GAA	TTT	GGT	TCA	TCT	TAT	CAG	TTG	TTG	48
Met	Asp	Ile	Asp	Pro	Tyr	Lys	Glu	Phe	Gly	Ser	Ser	Tyr	Gln	Leu	Leu	
1				5					10					15		
AAT	TTT	CTT	CCT	TTG	GAC	TTC	TTT	CCT	GAC	CTT	AAT	GCT	TTG	GTG	GAC	96
Asn	Phe	Leu	Pro	Leu	Asp	Phe	Phe	Pro	Asp	Leu	Asn	Ala	Leu	Val	Asp	
			20						25					30		

- 41 -

ACT	GCT	ACT	GCC	TTG	TAT	GAA	GAA	GAG	CTA	ACA	GGT	AGG	GAA	CAT	TGC	144
Thr	Ala	Thr	Ala	Leu	Tyr	Glu	Glu	Glu	Leu	Thr	Gly	Arg	Glu	His	Cys	
	35					40					45					
TCT	CCG	CAC	CAT	ACA	GCT	ATT	AGA	CAA	GCT	TTA	GTA	TGC	TGG	GAT	GAA	192
Ser	Pro	His	His	Thr	Ala	Ile	Arg	Gln	Ala	Leu	Val	Cys	Trp	Asp	Glu	
	50					55					60					
TTA	ACT	AAA	TTG	ATA	GCT	TGG	ATG	AGC	TCT	AAC	ATA	ACT	TCT	GAA	CAA	240
Leu	Thr	Lys	Leu	Ile	Ala	Trp	Met	Ser	Ser	Asn	Ile	Thr	Ser	Glu	Gln	
	65			70						75					80	
GTA	AGA	ACA	ATC	ATA	GTA	AAT	CAT	GTC	AAT	GAT	ACC	TGG	GGA	CTT	AAG	288
Val	Arg	Thr	Ile	Ile	Val	Asn	His	Val	Asn	Asp	Thr	Trp	Gly	Leu	Lys	
			85						90					95		
GTG	AGA	CAA	AGT	TTA	TGG	TTT	CAT	TTG	TCA	TGT	CTC	ACT	TTC	GGA	CAA	336
Val	Arg	Gln	Ser	Leu	Trp	Phe	His	Leu	Ser	Cys	Leu	Thr	Phe	Gly	Gln	
		100						105					110			
CAT	ACA	GTT	CAA	GAA	TTT	TTA	GTA	AGT	TTT	GTA	GTA	TGG	ATC	AGA	ACT	384
His	Thr	Val	Gln	Glu	Phe	Leu	Val	Ser	Phe	Val	Val	Trp	Ile	Arg	Thr	
		115				120						125				
CCA	GCT	CCA	TAT	AGA	CCT	CCT	AAT	GCA	CCC	ATT	CTC	TGG	ACT	CTT	CCG	432
Pro	Ala	Pro	Tyr	Arg	Pro	Pro	Asn	Ala	Pro	Ile	Leu	Ser	Thr	Leu	Pro	
	130					135					140					
GAA	CAT	ACA	GTC	ATT	AGA	AGA	GGA	GGT	GCA	AGA	GCT	TCT	AGG	TCC	CCC	480
Glu	His	Thr	Val	Ile	Arg	Arg	Gly	Gly	Ala	Arg	Ala	Ser	Arg	Ser	Pro	
	145			150					155						160	
AGA	AGA	CGC	ACT	CCC	TCT	CCT	CGC	AGG	AGA	AGA	TCC	CAA	AAT	TGG	CAG	528
Arg	Arg	Arg	Thr	Pro	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Gln	Asn	Ser	Gln	
				165					170					175		
TTC	CAA	ACT	TGC	AAA	CAC	TTG	CCA	ACC	TCC	TGT	CCA	CCA	ACT	TGC	AAT	576
Phe	Gln	Thr	Cys	Lys	His	Leu	Pro	Thr	Ser	Cys	Pro	Pro	Thr	Cys	Asn	
		180						185					190			
GGC	TTT	CGT	TGG	ATG	TAT	CTG	CGG	CGT	TTT	ATC	ATA	TAC	CTA	TTA	GTC	624
Gly	Phe	Arg	Trp	Met	Tyr	Leu	Arg	Arg	Phe	Ile	Ile	Tyr	Leu	Leu	Val	
	195			200								205				
CTG	CTG	CTG	TGC	CTC	ATC	TTC	TTG	TTG	GTT	CTC	CTG	GAC	TGG	AAA	GGT	672
Leu	Leu	Leu	Cys	Leu	Ile	Phe	Leu	Leu	Val	Leu	Leu	Asp	Trp	Lys	Gly	
	210					215					220					
TTA	ATA	CCT	GTC	TGT	CCT	CTT	CAA	CCC	ACA	ACA	GAA	ACA	ACA	GTC	AAT	720
Leu	Ile	Pro	Val	Cys	Pro	Leu	Gln	Pro	Thr	Thr	Glu	Thr	Thr	Val	Asn	
	225				230				235						240	
TGC	AGA	CAA	TGC	ACA	ATC	TCT	GCA	CAA	GAC	ATG	TAT	ACT	CCT	CCT	TAC	768
Cys	Arg	Gln	Cys	Thr	Ile	Ser	Ala	Gln	Asp	Met	Tyr	Thr	Pro	Pro	Tyr	
			245						250					255		
TGT	TGT	TGT	TTA	AAA	CCT	ACG	GCA	GGA	AAT	TGC	ACT	TGT	TGG	CCC	ATC	816
Cys	Cys	Cys	Leu	Lys	Pro	Thr	Ala	Gly	Asn	Cys	Thr	Cys	Trp	Pro	Ile	
			260					265					270			
CCT	TCA	TCA	TGG	GCT	TTA	GGA	AAT	TAC	CTA	TGG	GAG	TGG	GCC	TTA	GCT	864
Pro	Ser	Ser	Trp	Ala	Leu	Gly	Asn	Tyr	Leu	Trp	Glu	Trp	Ala	Leu	Ala	
		275					280					285				

- 42 -

CGT CTC TCT TGG CTC AAT TTA CTA GTG CCC TTG CTT CAA TGG TTA GGA	912
Arg Leu Ser Trp Leu Asn Leu Leu Val Pro Leu Leu Gln Trp Leu Gly	
290 295 300	
GGA ATT TCC CTC ATT GCG TGG TTT TTG CTT ATA TGG ATG ATT TGG TTT	960
Gly Ile Ser Leu Ile Ala Trp Phe Leu Leu Ile Trp Met Ile Trp Phe	
305 310 315 320	
TGG GGG CCC GCA CTT CTG AGC ATC TTA CCG CCA TTT ATT CCC ATA TTT	1008
Trp Gly Pro Ala Leu Leu Ser Ile Leu Pro Pro Phe Ile Pro Ile Phe	
325 330 335	
GTT CTG TTT TTC TTG ATT TGG GTA TAC ATT T GA	1041
Val Leu Phe Phe Leu Ile Trp Val Tyr Ile	
340 345	

## (2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 346 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ser Ser Tyr Gln Leu Leu	
1 5 10 15	
Asn Phe Leu Pro Leu Asp Phe Phe Pro Asp Leu Asn Ala Leu Val Asp	
20 25 30	
Thr Ala Thr Ala Leu Tyr Glu Glu Glu Leu Thr Gly Arg Glu His Cys	
35 40 45	
Ser Pro His His Thr Ala Ile Arg Gln Ala Leu Val Cys Trp Asp Glu	
50 55 60	
Leu Thr Lys Leu Ile Ala Trp Met Ser Ser Asn Ile Thr Ser Glu Gln	
65 70 75 80	
Val Arg Thr Ile Ile Val Asn His Val Asn Asp Thr Trp Gly Leu Lys	
85 90 95	
Val Arg Gln Ser Leu Trp Phe His Leu Ser Cys Leu Thr Phe Gly Gln	
100 105 110	
His Thr Val Gln Glu Phe Leu Val Ser Phe Val Val Trp Ile Arg Thr	
115 120 125	
Pro Ala Pro Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro	
130 135 140	
Glu His Thr Val Ile Arg Arg Gly Gly Ala Arg Ala Ser Arg Ser Pro	
145 150 155 160	
Arg Arg Arg Thr Pro Ser Pro Arg Arg Arg Arg Ser Gln Asn Ser Gln	
165 170 175	
Phe Gln Thr Cys Lys His Leu Pro Thr Ser Cys Pro Pro Thr Cys Asn	
180 185 190	

- 43 -

Gly Phe Arg Trp Met Tyr Leu Arg Arg Phe Ile Ile Tyr Leu Leu Val  
 195 200 205  
 Leu Leu Leu Cys Leu Ile Phe Leu Leu Val Leu Leu Asp Trp Lys Gly  
 210 215 220  
 Leu Ile Pro Val Cys Pro Leu Gln Pro Thr Thr Glu Thr Thr Val Asn  
 225 230 235 240  
 Cys Arg Gln Cys Thr Ile Ser Ala Gln Asp Met Tyr Thr Pro Pro Tyr  
 245 250 255  
 Cys Cys Cys Leu Lys Pro Thr Ala Gly Asn Cys Thr Cys Trp Pro Ile  
 260 265 270  
 Pro Ser Ser Trp Ala Leu Gly Asn Tyr Leu Trp Glu Trp Ala Leu Ala  
 275 280 285  
 Arg Leu Ser Trp Leu Asn Leu Leu Val Pro Leu Leu Gln Trp Leu Gly  
 290 295 300  
 Gly Ile Ser Leu Ile Ala Trp Phe Leu Leu Ile Trp Met Ile Trp Phe  
 305 310 315 320  
 Trp Gly Pro Ala Leu Leu Ser Ile Leu Pro Pro Phe Ile Pro Ile Phe  
 325 330 335  
 Val Leu Phe Phe Leu Ile Trp Val Tyr Ile  
 340 345

## (2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1056 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCT GGC AGG ATG GAC ATC GAC CCT TAT AAA GAA TTT GGA GCT ACT GTG	48
Pro Arg Arg Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val	
350 355 360	
GAG TTA CTC TCG TTT TTG CCT TCT GAC TTC TTT CCT TCA GTA CGA GAT	96
Glu Leu Leu Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp	
365 370 375	
CTT CTA GAT ACC GGC TCA GCT CTG TAT CGG GAA GCC TTA GAG TCT CCT	144
Leu Leu Asp Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro	
380 385 390	
GAG CAT TGT TCA CCT CAC CAT ACT GCA CTC AGG CAA GCA ATT CTT TGC	192
Glu His Cys Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys	
395 400 405 410	
TGG GGG GAA CTA ATG ACT CTA GCT ACC TGG GTG GGT GTT AAT TTG GAA	240
Trp Gly Glu Leu Met Thr Leu Ala Thr Trp Val Gly Val Asn Leu Glu	
415 420 425	

- 44 -

GAT	CCA	GCG	TCT	AGA	GAC	CTA	GTA	GTC	AGT	TAT	GTC	AAC	ACT	AAT	ATG	288
Asp	Pro	Ala	Ser	Arg	Asp	Leu	Val	Val	Ser	Tyr	Val	Asn	Thr	Asn	Met	
			430					435					440			
GGC	CTA	AAG	TTC	AGG	CAA	CTC	TTG	TGG	TTT	CAC	ATT	TCT	TGT	CTC	ACT	336
Gly	Leu	Lys	Phe	Arg	Gln	Leu	Leu	Trp	Phe	His	Ile	Ser	Cys	Leu	Thr	
		445					450					455				
TTT	GGA	AGA	GAA	ACA	GTT	ATA	GAG	TAT	TTG	GTG	TCT	TTC	GGA	GTG	TGG	384
Phe	Gly	Arg	Glu	Thr	Val	Ile	Glu	Tyr	Leu	Val	Ser	Phe	Gly	Val	Trp	
	460					465					470					
ATT	CGC	ACT	CCT	CCA	GCT	TAT	AGA	CCA	CCA	AAT	GCC	CCT	ATC	CTA	TCA	432
Ile	Arg	Thr	Pro	Pro	Ala	Tyr	Arg	Pro	Pro	Asn	Ala	Pro	Ile	Leu	Ser	
	475				480					485					490	
ACA	CTT	CCG	GAA	CAT	ACA	GTC	ATT	AGA	AGA	GGA	GGT	GCA	AGA	GCT	TCT	480
Thr	Leu	Pro	Glu	His	Thr	Val	Ile	Arg	Arg	Gly	Gly	Ala	Arg	Ala	Ser	
				495						500				505		
AGG	TCC	CCC	AGA	AGA	CGC	ACT	CCC	TCT	CCT	CGC	AGG	AGA	AGA	TCC	CAA	528
Arg	Ser	Pro	Arg	Arg	Arg	Thr	Pro	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Gln	
			510					515					520			
AAT	TCG	CAG	TCC	CCA	ACC	TCC	AAT	CAC	TCA	CCA	ACC	TCT	TGT	CCT	CCA	576
Asn	Ser	Gln	Ser	Pro	Thr	Ser	Asn	His	Ser	Pro	Thr	Ser	Cys	Pro	Pro	
		525					530					535				
ACT	TGT	CCT	GGT	TAT	CGC	TGG	ATG	TGT	CTG	CGG	CGT	TTT	ATC	ATC	TTC	624
Thr	Cys	Pro	Gly	Tyr	Arg	Trp	Met	Cys	Leu	Arg	Arg	Phe	Ile	Ile	Phe	
		540				545					550					
CTC	TTC	ATC	CTG	CTG	CTA	TGC	CTC	ATC	TTC	TTG	TTG	GTT	CTT	CTG	GAC	672
Leu	Phe	Ile	Leu	Leu	Leu	Cys	Leu	Ile	Phe	Leu	Leu	Val	Leu	Leu	Asp	
					560					565					570	
TAT	CAA	GGT	ATG	TTG	CCC	GTT	TGT	CCT	CTA	ATT	CCA	GGA	TCC	TCA	ACA	720
Tyr	Gln	Gly	Met	Leu	Pro	Val	Cys	Pro	Leu	Ile	Pro	Gly	Ser	Ser	Thr	
				575					580					585		
ACC	AGC	ACG	GGA	CCA	TGC	CGG	ACC	TGC	ATG	ACT	ACT	GCT	CAA	GGA	ACC	768
Thr	Ser	Thr	Gly	Pro	Cys	Arg	Thr	Cys	Met	Thr	Thr	Ala	Gln	Gly	Thr	
			590					595					600			
TCT	ATG	TAT	CCC	TCC	TGT	TGC	TGT	ACC	AAA	CCT	TCG	GAC	GGA	AAT	TGC	816
Ser	Met	Tyr	Pro	Ser	Cys	Cys	Cys	Thr	Lys	Pro	Ser	Asp	Gly	Asn	Cys	
		605				610						615				
ACC	TGT	ATT	CCC	ATC	CCA	TCA	TCC	TGG	GCT	TTC	GGA	AAA	TTC	CTA	TGG	864
Thr	Cys	Ile	Pro	Ile	Pro	Ser	Ser	Trp	Ala	Phe	Gly	Lys	Phe	Leu	Trp	
		620				625					630					
GAG	TGG	GCC	TCA	GCC	OGT	TTC	TCC	TGG	CTC	AGT	TTA	CTA	GTG	CCA	TTT	912
Glu	Trp	Ala	Ser	Ala	Arg	Phe	Ser	Trp	Leu	Ser	Leu	Leu	Val	Pro	Phe	
				640						645					650	
GTT	CAG	TGG	TTC	GTA	GGG	CTT	TCC	CCC	ACT	GTT	TGG	CTT	TCA	GTT	ATA	960
Val	Gln	Trp	Phe	Val	Gly	Leu	Ser	Pro	Thr	Val	Trp	Leu	Ser	Val	Ile	
				655					660					665		
TGG	ATG	ATG	TGG	TAT	TGG	GGG	CCA	AGT	CTG	TAC	AGC	ATC	TTG	AGT	CCC	1008
Trp	Met	Met	Trp	Tyr	Trp	Gly	Pro	Ser	Leu	Tyr	Ser	Ile	Leu	Ser	Pro	
			670					675					680			

- 45 -

TTT TTA CCG CTG TTA CCA ATT TTC TTT TGT CTT TGG GTA TAC ATT T 1054  
 Phe Leu Pro Leu Leu Pro Ile Phe Phe Cys Leu Trp Val Tyr Ile  
 685 690 695

AA 1056

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 351 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Pro Arg Arg Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val  
 1 5 10 15  
 Glu Leu Leu Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp  
 20 25 30  
 Leu Leu Asp Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro  
 35 40 45  
 Glu His Cys Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys  
 50 55 60  
 Trp Gly Glu Leu Met Thr Leu Ala Thr Trp Val Gly Val Asn Leu Glu  
 65 70 75 80  
 Asp Pro Ala Ser Arg Asp Leu Val Val Ser Tyr Val Asn Thr Asn Met  
 85 90 95  
 Gly Leu Lys Phe Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr  
 100 105 110  
 Phe Gly Arg Glu Thr Val Ile Glu Tyr Leu Val Ser Phe Gly Val Trp  
 115 120 125  
 Ile Arg Thr Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser  
 130 135 140  
 Thr Leu Pro Glu His Thr Val Ile Arg Arg Gly Gly Ala Arg Ala Ser  
 145 150 155 160  
 Arg Ser Pro Arg Arg Arg Thr Pro Ser Pro Arg Arg Arg Arg Ser Gln  
 165 170 175  
 Asn Ser Gln Ser Pro Thr Ser Asn His Ser Pro Thr Ser Cys Pro Pro  
 180 185 190  
 Thr Cys Pro Gly Tyr Arg Trp Met Cys Leu Arg Arg Phe Ile Ile Phe  
 195 200 205  
 Leu Phe Ile Leu Leu Leu Cys Leu Ile Phe Leu Leu Val Leu Leu Asp  
 210 215 220  
 Tyr Gln Gly Met Leu Pro Val Cys Pro Leu Ile Pro Gly Ser Ser Thr  
 225 230 235 240

Thr	Ser	Thr	Gly	Pro	Cys	Arg	Thr	Cys	Met	Thr	Thr	Ala	Gln	Gly	Thr
				245					250					255	
Ser	Met	Tyr	Pro	Ser	Cys	Cys	Cys	Thr	Lys	Pro	Ser	Asp	Gly	Asn	Cys
			260					265					270		
Thr	Cys	Ile	Pro	Ile	Pro	Ser	Ser	Trp	Ala	Phe	Gly	Lys	Phe	Leu	Trp
		275					280					285			
Glu	Trp	Ala	Ser	Ala	Arg	Phe	Ser	Trp	Leu	Ser	Leu	Leu	Val	Pro	Phe
	290					295					300				
Val	Gln	Trp	Phe	Val	Gly	Leu	Ser	Pro	Thr	Val	Trp	Leu	Ser	Val	Ile
					310					315					320
Trp	Met	Met	Trp	Tyr	Trp	Gly	Pro	Ser	Leu	Tyr	Ser	Ile	Leu	Ser	Pro
				325					330					335	
Phe	Leu	Pro	Leu	Leu	Pro	Ile	Phe	Phe	Cys	Leu	Trp	Val	Tyr	Ile	
			340					345					350		

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1194 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

17:39:34

- 47 -

GAA Glu 465	ACA Thr 465	GTT Val 465	ATA Ile 465	GAG Glu 465	TAT Tyr 465	TTG Leu 470	GTG Val 470	TCT Ser 470	TTC Phe 470	GGA Gly 475	GTG Val 475	TGG Trp 475	ATT Ile 475	CGC Arg 475	ACT Thr 475	384
CCT Pro 480	CCA Pro 480	GCT Ala 480	TAT Tyr 480	AGA Arg 485	CCA Pro 485	CCA Pro 485	AAT Asn 490	GCC Ala 490	CCT Pro 490	ATC Ile 490	CTA Leu 490	TCA Ser 490	ACA Thr 495	CTT Leu 495	CCG Pro 495	432
GAG Glu 500	ACT Thr 500	ACT Thr 500	GTT Val 500	GTT Val 500	AGA Arg 500	CGA Arg 500	CGA Arg 505	GGC Gly 505	AGG Ser 505	TCC Ser 505	CCT Pro 505	AGA Arg 510	AGA Arg 510	AGA Arg 510	ACT Thr 510	480
CCC Pro 515	TGG Ser 515	CCT Pro 515	CGC Arg 515	AGA Arg 515	CGA Arg 515	AGG Arg 515	TCT Ser 520	CAA Gln 520	TGG Ser 520	CCG Pro 520	CGT Arg 520	CGC Arg 525	AGA Arg 525	AGA Arg 525	TCT Ser 525	528
CAA Gln 530	TCT Ser 530	CGG Arg 530	CTA Leu 530	GGA Gly 530	CCC Pro 530	CTT Leu 535	CTC Leu 535	GTG Val 535	TTA Leu 535	CAG Gln 540	GCG Ala 540	GGG Gly 540	TTT Phe 540	TTC Phe 540	TTG Leu 540	576
TTG Leu 545	ACA Thr 545	AGA Arg 545	ATC Ile 545	CTC Leu 545	ACA Thr 545	ATA Ile 550	CCG Pro 550	CAG Gln 550	AGT Ser 550	CTA Leu 555	GAC Asp 555	TGG Ser 555	TGG Trp 555	TGG Trp 555	ACT Thr 555	624
TCT Ser 560	CTC Leu 560	AAT Asn 560	TTT Phe 560	CTA Leu 565	GGG Gly 565	GGA Gly 565	ACT Thr 565	ACC Thr 565	GTG Val 570	TGT Cys 570	CTT Leu 570	GGC Gly 570	CAA Gln 575	AAT Asn 575	TGG Ser 575	672
CAG Gln 580	TCC Ser 580	CCA Pro 580	ACC Thr 580	TCC Ser 580	AAT Asn 580	CAC His 580	TCA Ser 585	CCA Pro 585	ACC Thr 585	TCT Ser 585	TGT Cys 585	CCT Pro 590	CCA Pro 590	ACT Thr 590	TGT Cys 590	720
CCT Pro 595	GGT Gly 595	TAT Tyr 595	CGC Arg 595	TGG Trp 595	ATG Met 595	TGT Cys 595	CTG Leu 600	CGG Arg 600	CGT Arg 600	TTT Phe 605	ATC Ile 605	ATC Ile 605	TTC Phe 605	CTC Leu 605	TTC Phe 605	768
ATC Ile 610	CTG Leu 610	CTG Leu 610	CTA Leu 610	TGC Cys 610	CTC Leu 610	ATC Ile 615	TTC Phe 615	TTG Leu 615	TTG Leu 615	GTT Val 620	CTT Leu 620	CTG Leu 620	GAC Asp 620	TAT Tyr 620	CAA Gln 620	816
GGT Gly 625	ATG Met 625	TTG Leu 625	CCC Pro 625	GTT Val 625	TGT Cys 625	CCT Pro 630	CTA Leu 630	ATT Ile 630	CCA Pro 630	GGA Gly 635	TCC Ser 635	TCA Ser 635	ACA Thr 635	ACC Thr 635	AGC Ser 635	864
ACG Thr 640	GGA Gly 640	CCA Pro 640	TGC Cys 640	CGG Arg 645	ACC Thr 645	TGC Cys 645	ATG Met 645	ACT Thr 645	ACT Thr 645	GCT Ala 650	CAA Gln 650	GGA Gly 650	ACC Thr 650	TCT Ser 655	ATG Met 655	912
TAT Tyr 660	CCC Pro 660	TCC Ser 660	TGT Cys 660	TGC Cys 660	TGT Cys 660	ACC Thr 660	AAA Lys 665	CCT Pro 665	TGG Ser 665	GAC Asp 665	GGA Gly 665	AAT Asn 665	TGC Cys 670	ACC Thr 670	TGT Cys 670	960
ATT Ile 675	CCC Pro 675	ATC Ile 675	CCA Pro 675	TCA Ser 675	TCC Ser 675	TGG Trp 675	GCT Ala 680	TTC Phe 680	GGA Gly 680	AAA Lys 685	TTC Phe 685	CTA Leu 685	TGG Trp 685	GAG Glu 685	TGG Trp 685	1008
GCC Ala 690	TCA Ser 690	GCC Ala 690	CGT Arg 690	TTC Phe 690	TCC Ser 690	TGG Trp 695	CTC Leu 695	AGT Ser 695	TTA Leu 695	CTA Leu 700	GTG Val 700	CCA Pro 700	TTT Phe 700	GTT Val 700	CAG Gln 700	1056
TGG Trp 705	TTC Phe 705	GTA Val 705	GGG Gly 705	CTT Leu 705	TCC Ser 705	CCC Pro 710	ACT Thr 710	GTT Val 710	TGG Trp 710	CTT Leu 715	TCA Ser 715	GTT Val 715	ATA Ile 715	TGG Trp 715	ATG Met 715	1104

- 48 -

ATG TGG TAT TGG GGG CCA AGT CTG TAC AGC ATC TTG AGT CCC TTT TTA	1152
Met Trp Tyr Trp Gly Pro Ser Leu Tyr Ser Ile Leu Ser Pro Phe Leu	
720 725 730 735	
CCG CTG TTA CCA ATT TTC TTT TGT CTT TGG GTA TAC ATT T AA	1194
Pro Leu Leu Pro Ile Phe Phe Cys Leu Trp Val Tyr Ile	
740 745	

## (2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 397 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met	Asp	Ile	Asp	Pro	Tyr	Lys	Glu	Phe	Gly	Ala	Thr	Val	Glu	Leu	Leu	
1				5					10					15		
Ser	Phe	Leu	Pro	Ser	Asp	Phe	Phe	Pro	Ser	Val	Arg	Asp	Leu	Leu	Asp	
			20					25					30			
Thr	Ala	Ser	Ala	Leu	Tyr	Arg	Glu	Ala	Leu	Glu	Ser	Pro	Glu	His	Cys	
			35				40					45				
Ser	Pro	His	His	Thr	Ala	Leu	Arg	Gln	Ala	Ile	Leu	Cys	Trp	Gly	Glu	
	50					55					60					
Leu	Met	Thr	Leu	Ala	Thr	Trp	Val	Gly	Val	Asn	Leu	Glu	Asp	Pro	Ala	
	65				70					75					80	
Ser	Arg	Asp	Leu	Val	Val	Ser	Tyr	Val	Asn	Thr	Asn	Met	Gly	Leu	Lys	
			85						90					95		
Phe	Arg	Gln	Leu	Leu	Trp	Phe	His	Ile	Ser	Cys	Leu	Thr	Phe	Gly	Arg	
			100				105						110			
Glu	Thr	Val	Ile	Glu	Tyr	Leu	Val	Ser	Phe	Gly	Val	Trp	Ile	Arg	Thr	
			115				120					125				
Pro	Pro	Ala	Tyr	Arg	Pro	Pro	Asn	Ala	Pro	Ile	Leu	Ser	Thr	Leu	Pro	
			130			135					140					
Glu	Thr	Thr	Val	Val	Arg	Arg	Arg	Gly	Arg	Ser	Pro	Arg	Arg	Arg	Thr	
	145				150					155					160	
Pro	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Gln	Ser	Pro	Arg	Arg	Arg	Arg	Ser	
				165					170					175		
Gln	Ser	Arg	Leu	Gly	Pro	Leu	Leu	Val	Leu	Gln	Ala	Gly	Phe	Phe	Leu	
			180				185						190			
Leu	Thr	Arg	Ile	Leu	Thr	Ile	Pro	Gln	Ser	Leu	Asp	Ser	Trp	Trp	Thr	
			195				200					205				
Ser	Leu	Asn	Phe	Leu	Gly	Gly	Thr	Thr	Val	Cys	Leu	Gly	Gln	Asn	Ser	
						215					220					

- 49 -

Gln Ser Pro Thr Ser Asn His Ser Pro Thr Ser Cys Pro Pro Thr Cys  
 225 230 235 240  
 Pro Gly Tyr Arg Trp Met Cys Leu Arg Arg Phe Ile Ile Phe Leu Phe  
 245 250 255  
 Ile Leu Leu Leu Cys Leu Ile Phe Leu Leu Val Leu Leu Asp Tyr Gln  
 260 265 270  
 Gly Met Leu Pro Val Cys Pro Leu Ile Pro Gly Ser Ser Thr Thr Ser  
 275 280 285  
 Thr Gly Pro Cys Arg Thr Cys Met Thr Thr Ala Gln Gly Thr Ser Met  
 290 295 300  
 Tyr Pro Ser Cys Cys Cys Thr Lys Pro Ser Asp Gly Asn Cys Thr Cys  
 305 310 315 320  
 Ile Pro Ile Pro Ser Ser Trp Ala Phe Gly Lys Phe Leu Trp Glu Trp  
 325 330 335  
 Ala Ser Ala Arg Phe Ser Trp Leu Ser Leu Leu Val Pro Phe Val Gln  
 340 345 350  
 Trp Phe Val Gly Leu Ser Pro Thr Val Trp Leu Ser Val Ile Trp Met  
 355 360 365  
 Met Trp Tyr Trp Gly Pro Ser Leu Tyr Ser Ile Leu Ser Pro Phe Leu  
 370 375 380  
 Pro Leu Leu Pro Ile Phe Phe Cys Leu Trp Val Tyr Ile  
 385 390 395

## (2) INFORMATION FOR SEQ ID NO:7:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 870 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (11) MOLECULE TYPE: DNA

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATG GAC ATC GAC CCT TAT AAA GAA TTT GGA GCT ACT GTG GAG TTA CTC	48
Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu	
400 405 410	
TCG TTT TTG CCT TCT GAC TTC TTT CCT TCA GTA CGA GAT CTT CTA GAT	96
Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp	
415 420 425	
ACC GCC TCA GCT CTG TAT CGG GAA GCC TTA GAG TCT CCT GAG CAT TGT	144
Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys	
430 435 440 445	
TCA CCT CAC CAT ACT GCA CTC AGG CAA GCA ATT CTT TGC TCG GGG GAA	192
Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu	
450 455 460	

- 50 -

CTA	ATG	ACT	CTA	GCT	ACC	TGG	GTG	GGT	GTT	AAT	TTG	GAA	GAT	CCA	CCG	240
Leu	Met	Thr	Leu	Ala	Thr	Trp	Val	Gly	Val	Asn	Leu	Glu	Asp	Pro	Ala	
			465					470					475			
TCT	AGA	GAC	CTA	GTA	GTC	ACT	TAT	GTC	AAC	ACT	AAT	ATG	GGC	CTA	AAG	288
Ser	Arg	Asp	Leu	Val	Val	Ser	Tyr	Val	Asn	Thr	Asn	Met	Gly	Leu	Lys	
		480					485					490				
TTC	AGG	CAA	CTC	TTG	TGG	TTT	CAC	ATT	TCT	TGT	CTC	ACT	TTT	GGA	AGA	336
Phe	Arg	Gln	Leu	Leu	Trp	Phe	His	Ile	Ser	Cys	Leu	Thr	Phe	Gly	Arg	
	495					500					505					
GAA	ACA	GTT	ATA	GAG	TAT	TTG	GTG	TCT	TTC	GGA	GTG	TGG	ATT	CGC	ACT	384
Glu	Thr	Val	Ile	Glu	Tyr	Leu	Val	Ser	Phe	Gly	Val	Trp	Ile	Arg	Thr	
510				515						520				525		
CCT	CCA	GCT	TAT	AGA	CCA	CCA	AAT	GCC	CCT	ATC	CTA	TCA	ACA	CTT	CCG	432
Pro	Pro	Ala	Tyr	Arg	Pro	Pro	Asn	Ala	Pro	Ile	Leu	Ser	Thr	Pro	Pro	
				530				535						540		
GAG	ACT	ACT	GTT	GTT	AGA	CGA	CGA	GGC	AGG	TCC	CCT	AGA	AGA	AGA	ACT	480
Glu	Thr	Thr	Val	Val	Arg	Arg	Arg	Gly	Arg	Ser	Pro	Arg	Arg	Arg	Thr	
			545					550					555			
CCC	TCG	CCT	CGC	AGA	CGA	AGG	TCT	CAA	TCG	CCG	CGT	CGC	AGA	AGA	TCC	528
Pro	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Gln	Ser	Pro	Arg	Arg	Arg	Arg	Ser	
		560					565					570				
TCA	ACA	ACC	AGC	ACG	GGA	CCA	TGC	CGG	ACC	TGC	ATG	ACT	ACT	GCT	CAA	576
Ser	Thr	Thr	Ser	Thr	Gly	Pro	Cys	Arg	Thr	Cys	Met	Thr	Thr	Ala	Gln	
		575				580					585					
GGA	ACC	TCT	ATG	TAT	CCC	TCC	TGT	TGC	TGT	ACC	AAA	CCT	TCG	GAC	GGA	624
Gly	Thr	Ser	Met	Tyr	Pro	Ser	Cys	Cys	Cys	Thr	Lys	Pro	Ser	Asp	Gly	
590					595					600				605		
AAT	TGC	ACC	TGT	ATT	CCC	ATC	CCA	TCA	TCC	TGG	GCT	TTC	GGA	AAA	TTC	672
Asn	Cys	Thr	Cys	Ile	Pro	Ile	Pro	Ser	Ser	Trp	Ala	Phe	Gly	Lys	Phe	
				610				615						620		
CTA	TGG	GAG	TGG	GCC	TCA	GCC	CGT	TTC	TCC	TGG	CTC	AGT	TTA	CTA	GTG	720
Leu	Trp	Glu	Trp	Ala	Ser	Ala	Arg	Phe	Ser	Trp	Leu	Ser	Leu	Leu	Val	
			625					630					635			
CCA	TTT	GTT	CAG	TGG	TTC	GTA	GGG	CTT	TCC	CCC	ACT	GTT	TGG	CTT	TCA	768
Pro	Phe	Val	Gln	Trp	Phe	Val	Gly	Leu	Ser	Pro	Thr	Val	Trp	Leu	Ser	
		640					645					650				
GTT	ATA	TGG	ATG	ATG	TGG	TAT	TGG	GGG	CCA	AGT	CTG	TAC	AGC	ATC	TTG	816
Val	Ile	Trp	Met	Met	Trp	Tyr	Trp	Gly	Pro	Ser	Leu	Tyr	Ser	Ile	Leu	
		655				660					665					
AGT	CCC	TTT	TTA	CCG	CTG	TTA	CCA	ATT	TTC	TTT	TGT	CTT	TGG	GTA	TAC	864
Ser	Pro	Phe	Leu	Pro	Leu	Leu	Pro	Ile	Phe	Cys	Leu	Trp	Val	Tyr		
670					675					680				685		
ATT	T	AA														870
Ile																

(2) INFORMATION FOR SEQ ID NO:8:

- 51 -

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 289 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```

Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu
 1           5           10           15
Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp
          20           25           30
Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys
          35           40           45
Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu
          50           55           60
Leu Met Thr Leu Ala Thr Trp Val Gly Val Asn Leu Glu Asp Pro Ala
          65           70           75           80
Ser Arg Asp Leu Val Val Ser Tyr Val Asn Thr Asn Met Gly Leu Lys
          85           90           95
Phe Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg
          100          105          110
Glu Thr Val Ile Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr
          115          120          125
Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro
          130          135          140
Glu Thr Thr Val Val Arg Arg Arg Gly Arg Ser Pro Arg Arg Arg Thr
          145          150          155          160
Pro Ser Pro Arg Arg Arg Arg Ser Gln Ser Pro Arg Arg Arg Arg Ser
          165          170          175
Ser Thr Thr Ser Thr Gly Pro Cys Arg Thr Cys Met Thr Thr Ala Gln
          180          185          190
Gly Thr Ser Met Tyr Pro Ser Cys Cys Cys Thr Lys Pro Ser Asp Gly
          195          200          205
Asn Cys Thr Cys Ile Pro Ile Pro Ser Ser Trp Ala Phe Gly Lys Phe
          210          215          220
Leu Trp Glu Trp Ala Ser Ala Arg Phe Ser Trp Leu Ser Leu Leu Val
          225          230          235          240
Pro Phe Val Gln Trp Phe Val Gly Leu Ser Pro Thr Val Trp Leu Ser
          245          250          255
Val Ile Trp Met Met Trp Tyr Trp Gly Pro Ser Leu Tyr Ser Ile Leu
          260          265          270
Ser Pro Phe Leu Pro Leu Leu Pro Ile Phe Phe Cys Leu Trp Val Tyr
          275          280          285

```

- 52 -

ile

## (2) INFORMATION FOR SEQ ID NO:9:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1263 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ATG	GAT	ATC	AAT	GCT	TCT	AGA	GCC	TTA	GCC	AAT	GTG	TAT	GAT	CTA	CCA	48
Met	Asp	Ile	Asn	Ala	Ser	Arg	Ala	Leu	Ala	Asn	Val	Tyr	Asp	Leu	Pro	
290					295				300						305	
GAT	GAT	TTC	TTT	CCA	AAA	ATA	GAT	GAT	CTT	GTT	AGA	GAT	GCT	AAA	GAC	96
Asp	Asp	Phe	Phe	Pro	Lys	Ile	Asp	Asp	Leu	Val	Arg	Asp	Ala	Lys	Asp	
				310					315					320		
GCT	TTA	GAG	CCT	TAT	TGG	AAA	TCA	GAT	TCA	ATA	AAG	AAA	CAT	GTT	TTG	144
Ala	Leu	Glu	Pro	Tyr	Trp	Lys	Ser	Asp	Ser	Ile	Lys	Lys	His	Val	Leu	
			325					330					335			
ATT	GCA	ACT	CAC	TTT	GTG	GAT	CTT	ATT	GAA	GAC	TTC	TGG	CAG	ACT	ACA	192
Ile	Ala	Thr	His	Phe	Val	Asp	Leu	Ile	Glu	Asp	Phe	Trp	Gln	Thr	Thr	
		340				345					350					
CAG	GGC	ATG	CAT	GAA	ATA	GCC	GAA	TCA	TTA	AGA	GCT	GTT	ATA	CCT	CCC	240
Gln	Gly	Met	His	Glu	Ile	Ala	Glu	Ser	Leu	Arg	Ala	Val	Ile	Pro	Pro	
	355				360				365							
ACT	ACT	ACT	CCT	GTT	CCA	CCG	GGT	TAT	CTT	ATT	CAG	CAC	GAA	GAA	GCT	288
Thr	Thr	Thr	Pro	Val	Pro	Pro	Gly	Tyr	Leu	Ile	Gln	His	Glu	Glu	Ala	
			370		375				380					385		
GAA	GAG	ATA	CCT	TTG	GGA	GAT	TTA	TTT	AAA	CAC	CAA	GAA	GAA	AGG	ATA	336
Glu	Glu	Ile	Pro	Leu	Gly	Asp	Leu	Phe	Lys	His	Gln	Glu	Glu	Arg	Ile	
			390				395							400		
GTG	AGT	TTC	CAA	CCC	GAC	TAT	CCG	ATT	ACG	GCT	AGA	ATT	CAT	GCT	CAT	384
Val	Ser	Phe	Gln	Pro	Asp	Tyr	Pro	Ile	Thr	Ala	Arg	Ile	His	Ala	His	
		405				410						415				
TTG	AAA	GCT	TAT	GCA	AAA	ATT	AAC	GAG	GAA	TCA	CTG	GAT	AGG	GCT	AGG	432
Leu	Lys	Ala	Tyr	Ala	Lys	Ile	Asn	Glu	Glu	Ser	Leu	Asp	Arg	Ala	Arg	
		420				425					430					
AGA	TTG	CTT	TGG	TGG	CAT	TAC	AAC	TGT	TTA	CTG	TGG	GGA	GAA	GCT	CAA	480
Arg	Leu	Leu	Trp	Trp	His	Tyr	Asn	Cys	Leu	Leu	Trp	Gly	Glu	Ala	Gln	
		435			440						445					
GTT	ACT	AAC	TAT	ATT	TCT	CGC	TTG	CGT	ACT	TGG	TTG	TCA	ACT	CCT	GAG	528
Val	Thr	Asn	Tyr	Ile	Ser	Arg	Leu	Arg	Thr	Trp	Leu	Ser	Thr	Pro	Glu	
		450			455				460						465	

- 53 -

AAA TAT AGA GGT AGA GAT GCC CCG ACC ATT GAA GCA ATC ACT AGA CCA Lys Tyr Arg Gly Arg Asp Ala Pro Thr Ile Glu Ala Ile Thr Arg Pro	576
470 475 480	
ATC CAG GTG GCT CAG GGA GGC CGA AAA ACA ACT ACG GGT ACT AGA AAA Ile Gln Val Ala Gln Gly Gly Arg Lys Thr Thr Thr Gly Thr Arg Lys	624
485 490 495	
CCT CGT GGA CTC GAA CCT AGA AGA AGA AAA GTT AAA ACC ACA GTT GTC Pro Arg Gly Leu Glu Pro Arg Arg Arg Lys Val Lys Thr Thr Val Val	672
500 505 510	
TAT GGG AGA AGA CGT TCA AAG TCC CGG GGA AGG AGA GCC CCT ACA CCC Tyr Gly Arg Arg Arg Ser Lys Ser Arg Gly Arg Arg Ala Pro Thr Pro	720
515 520 525	
CAA CGT GCG GGC TCC CCT CTC CCA CGT AGT TCG AGC AGC CAC CAT AGA Gln Arg Ala Gly Ser Pro Leu Pro Arg Ser Ser Ser Ser His His Arg	768
530 535 540 545	
TCC TTC GGG GGA ATA CTA GCT GCC CTA ATC GGA TTA CTG GTA AGC TTT Ser Phe Gly Gly Ile Leu Ala Gly Leu Ile Gly Leu Leu Val Ser Phe	816
550 555 560	
TTC TTG TTG ATA AAA ATT CTA GAA ATA CTG AGG AGG CTA GAT TGG TGG Phe Leu Leu Ile Lys Ile Leu Glu Ile Leu Arg Arg Leu Asp Trp Trp	864
565 570 575	
TGG ATT TCT CTC AGT TCT CCA AAG GGA AAA ATG CAA TGC GCT TTC CAA Trp Ile Ser Leu Ser Ser Pro Lys Gly Lys Met Gln Cys Ala Phe Gln	912
580 585 590	
GAT ACT GGA GCC CAA ATC TCT CCA CAT TAC GTC GGA TCT TGC CCG TGG Asp Thr Gly Ala Gln Ile Ser Pro His Tyr Val Gly Ser Cys Pro Trp	960
595 600 605	
GGA TGC CCA GGA TTT CTT TGG ACC TAT CTC AGG CTT TTT ATC ATC TTC Gly Cys Pro Gly Phe Leu Trp Thr Tyr Leu Arg Leu Phe Ile Ile Phe	1008
610 615 620 625	
CTC TTA ATC CTG CTA GTA GCA GCA GGC TTG CTG TAT CTG ACG GAC AAC Leu Leu Ile Leu Leu Val Ala Ala Gly Leu Leu Tyr Leu Thr Asp Asn	1056
630 635 640	
GGG TCT ACT ATT TTA GGA AAG CTC CAA TGG GCG TCG GTC TCA GCC CTT Gly Ser Thr Ile Leu Gly Lys Leu Gln Trp Ala Ser Val Ser Ala Leu	1104
645 650 655	
TTC TCC TCC ATC TCT TCA CTA CTG CCC TCG GAT CCG AAA TCT CTC GTC Phe Ser Ser Ile Ser Ser Leu Leu Pro Ser Asp Pro Lys Ser Leu Val	1152
660 665 670	
GCT TTA ACG TTT GGA CTT TCA CTT ATA TGG ATG ACT TCC TCC TCT GCC Ala Leu Thr Phe Gly Leu Ser Leu Ile Trp Met Thr Ser Ser Ser Ala	1200
675 680 685	
ACC CAA ACG CTC GTC ACC TTA ACG CAA TTA GCC ACG CTG TCT GCT CTT Thr Gln Thr Leu Val Thr Leu Thr Gln Leu Ala Thr Leu Ser Ala Leu	1248
690 695 700 705	
TTT TAC AAG AGC T AG Phe Tyr Lys Ser	1263

- 54 -

## (2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 420 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```

Met Asp Ile Asn Ala Ser Arg Ala Leu Ala Asn Val Tyr Asp Leu Pro
 1           5           10           15
Asp Asp Phe Phe Pro Lys Ile Asp Asp Leu Val Arg Asp Ala Lys Asp
          20           25           30
Ala Leu Glu Pro Tyr Trp Lys Ser Asp Ser Ile Lys Lys His Val Leu
          35           40           45
Ile Ala Thr His Phe Val Asp Leu Ile Glu Asp Phe Trp Gln Thr Thr
          50           55           60
Gln Gly Met His Glu Ile Ala Glu Ser Leu Arg Ala Val Ile Pro Pro
          65           70           75           80
Thr Thr Thr Pro Val Pro Pro Gly Tyr Leu Ile Gln His Glu Glu Ala
          85           90           95
Glu Glu Ile Pro Leu Gly Asp Leu Phe Lys His Gln Glu Glu Arg Ile
          100          105          110
Val Ser Phe Gln Pro Asp Tyr Pro Ile Thr Ala Arg Ile His Ala His
          115          120          125
Leu Lys Ala Tyr Ala Lys Ile Asn Glu Glu Ser Leu Asp Arg Ala Arg
          130          135          140
Arg Leu Leu Trp Trp His Tyr Asn Cys Leu Leu Trp Gly Glu Ala Gln
          145          150          155          160
Val Thr Asn Tyr Ile Ser Arg Leu Arg Thr Trp Leu Ser Thr Pro Glu
          165          170          175
Lys Tyr Arg Gly Arg Asp Ala Pro Thr Ile Glu Ala Ile Thr Arg Pro
          180          185          190
Ile Gln Val Ala Gln Gly Gly Arg Lys Thr Thr Thr Gly Thr Arg Lys
          195          200          205
Pro Arg Gly Leu Glu Pro Arg Arg Arg Lys Val Lys Thr Thr Val Val
          210          215          220
Tyr Gly Arg Arg Arg Ser Lys Ser Arg Gly Arg Arg Ala Pro Thr Pro
          225          230          235          240
Gln Arg Ala Gly Ser Pro Leu Pro Arg Ser Ser Ser Ser His His Arg
          245          250          255
Ser Phe Gly Gly Ile Leu Ala Gly Leu Ile Gly Leu Leu Val Ser Phe
          260          265          270

```

- 55 -

Phe Leu Leu Ile Lys Ile Leu Glu Ile Leu Arg Arg Leu Asp Trp Trp  
 275 280 285  
 Trp Ile Ser Leu Ser Ser Pro Lys Gly Lys Met Gln Cys Ala Phe Gln  
 290 295 300  
 Asp Thr Gly Ala Gln Ile Ser Pro His Tyr Val Gly Ser Cys Pro Trp  
 305 310 315 320  
 Gly Cys Pro Gly Phe Leu Trp Thr Tyr Leu Arg Leu Phe Ile Ile Phe  
 325 330 335  
 Leu Leu Ile Leu Leu Val Ala Ala Gly Leu Leu Tyr Leu Thr Asp Asn  
 340 345 350  
 Gly Ser Thr Ile Leu Gly Lys Leu Gln Trp Ala Ser Val Ser Ala Leu  
 355 360 365  
 Phe Ser Ser Ile Ser Ser Leu Leu Pro Ser Asp Pro Lys Ser Leu Val  
 370 375 380  
 Ala Leu Thr Phe Gly Leu Ser Leu Ile Trp Met Thr Ser Ser Ser Ala  
 385 390 395 400  
 Thr Gln Thr Leu Val Thr Leu Thr Gln Leu Ala Thr Leu Ser Ala Leu  
 405 410 415  
 Phe Tyr Lys Ser  
 420

## (2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 552 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ATG GAC ATC GAC CCT TAT AAA GAA TTT GGA GCT ACT GTC GAG TTA CTC	48
Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu	
1 5 10 15	
TCG TTT TTG CCT TCT GAC TTC TTT CCT TCA GTA CGA GAT CTT CTA GAT	96
Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp	
20 25 30	
ACC GCC TCA GCT CTG TAT CGG GAA GCC TTA GAG TCT CCT GAG CAT TGT	144
Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys	
35 40 45	
TCA CCT CAC CAT ACT GCA CTC AGG CAA GCA ATT CTT TGC TGG GGG GAA	192
Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu	
50 55 60	
CTA ATG ACT CTA GCT ACC TGG GTG GGT GTT AAT TTG GAA GAT CCA GCG	240
Leu Met Thr Leu Ala Thr Trp Val Gly Val Asn Leu Glu Asp Pro Ala	
65 70 75 80	

- 56 -

TCT AGA GAC CTA GTA GTC AGT TAT GTC AAC ACT AAT ATG GGC CTA AAG	288
Ser Arg Asp Leu Val Ser Tyr Val Asn Thr Asn Met Gly Leu Lys	
85 90 95	
TTC AGG CAA CTC TTG TGG TTT CAC ATT TCT TGT CTC ACT TTT GGA ACA	336
Phe Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Thr	
100 105 110	
GAA ACA GTT ATA GAG TAT TTG GTG TCT TTC GGA GTG TGG ATT CGC ACT	384
Glu Thr Val Ile Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr	
115 120 125	
CCT CCA GCT TAT AGA CCA CCA AAT GCC CCT ATC CTA TCA ACA CTT CCG	432
Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro	
130 135 140	
GAG ACT ACT GTT GTT AGA CGA CCA GGC AGG TCC CCT AGA AGA AGA ACT	480
Glu Thr Thr Val Val Arg Arg Pro Gly Arg Ser Pro Arg Arg Arg Thr	
145 150 155 160	
CCC TCG CCT CGC AGA CGA AGG TCT CAA TCG CCC CGT CGC AGA AGA TCT	528
Pro Ser Pro Arg Arg Arg Arg Ser Gln Ser Pro Arg Arg Arg Arg Ser	
165 170 175	
CAA TCT CCG GAA TCT CAA TGT TAG	552
Gln Ser Arg Glu Ser Gln Cys	
180	

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 183 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu	
1 5 10 15	
Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp	
20 25 30	
Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys	
35 40 45	
Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu	
50 55 60	
Leu Met Thr Leu Ala Thr Trp Val Gly Val Asn Leu Glu Asp Pro Ala	
65 70 75 80	
Ser Arg Asp Leu Val Val Ser Tyr Val Asn Thr Asn Met Gly Leu Lys	
85 90 95	
Phe Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Thr	
100 105 110	
Glu Thr Val Ile Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr	
115 120 125	

- 57 -

Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro  
 130 135 140

Glu Thr Thr Val Val Arg Arg Pro Gly Arg Ser Pro Arg Arg Arg Thr  
 145 150 155 160

Pro Ser Pro Arg Arg Arg Arg Ser Gln Ser Pro Arg Arg Arg Arg Ser  
 165 170 175

Gln Ser Arg Glu Ser Gln Cys  
 180

## (2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 681 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

ATC	GAG	AAC	ATC	ACA	TCA	GGA	TTC	CTA	GGA	CCC	CTT	CTC	CTG	TTA	CAG	48
Met	Glu	Asn	Ile	Thr	Ser	Gly	Phe	Leu	Gly	Pro	Leu	Leu	Val	Leu	Gln	
1				5					10					15		
GGG	GGG	TTT	TTC	TTG	TTG	ACA	AGA	ATC	CTC	ACA	ATA	CCC	CAG	AGT	CTA	96
Ala	Gly	Phe	Phe	Leu	Leu	Thr	Arg	Ile	Leu	Thr	Ile	Pro	Gln	Ser	Leu	
			20					25					30			
GAC	TCG	TGG	TGG	ACT	TCT	CTC	AAT	TTT	CTA	GGG	GGA	ACT	ACC	GTG	TGT	144
Asp	Ser	Trp	Trp	Thr	Ser	Leu	Asn	Phe	Leu	Gly	Gly	Thr	Thr	Val	Cys	
		35					40					45				
CTT	GGC	CAA	AAT	TCG	CAG	TCC	CCA	ACC	TCC	AAT	CAC	TCA	CCA	ACC	TCT	192
Leu	Gly	Gln	Asn	Ser	Gln	Ser	Pro	Thr	Ser	Asn	His	Ser	Pro	Thr	Ser	
		50				55					60					
TGT	CCT	CCA	ACT	TGT	CCT	GGT	TAT	CGC	TGC	ATG	TGT	CTG	CGG	CGT	TTT	240
Cys	Pro	Pro	Thr	Cys	Pro	Gly	Tyr	Arg	Trp	Met	Cys	Leu	Arg	Arg	Phe	
		65			70				75					80		
ATC	ATC	TTC	CTC	TTC	ATC	CTG	CTG	CTA	TGC	CTC	ATC	TTC	TTG	TTG	GTT	288
Ile	Ile	Phe	Leu	Phe	Ile	Leu	Leu	Leu	Cys	Leu	Ile	Phe	Leu	Leu	Val	
				85					90					95		
CTT	CTG	GAC	TAT	CAA	GGT	ATG	TTG	CCC	GTT	TGT	CCT	CTA	ATT	CCA	GCA	336
Leu	Leu	Asp	Tyr	Gln	Gly	Met	Leu	Pro	Val	Cys	Pro	Leu	Ile	Pro	Gly	
			100					105					110			
TCC	TCA	ACA	ACC	AGC	ACG	GGA	CCA	TGC	CGG	ACC	TGC	ATG	ACT	ACT	GCT	384
Ser	Ser	Thr	Thr	Ser	Thr	Gly	Pro	Cys	Arg	Thr	Cys	Met	Thr	Thr	Ala	
			115				120					125				
CAA	GGA	ACC	TCT	ATG	TAT	CCC	TCC	TGT	TGC	TGT	ACC	AAA	CCT	TCG	GAC	432
Gln	Gly	Thr	Ser	Met	Tyr	Pro	Ser	Cys	Cys	Cys	Thr	Lys	Pro	Ser	Asp	
		130				135					140					
GGA	AAT	TGC	ACC	TGT	ATT	CCC	ATC	CCA	TCA	TCC	TGG	GCT	TTC	GGA	AAA	480
Gly	Asn	Cys	Thr	Cys	Ile	Pro	Ile	Pro	Ser	Ser	Trp	Ala	Phe	Gly	Lys	
145					150					155				160		

- 58 -

TTC CTA TGG GAG TGG GCC TCA GCC CCT TTC TCC TGG CTC AGT TTA CTA	528
Phe Leu Trp Glu Trp Ala Ser Ala Pro Phe Ser Trp Leu Ser Leu Leu	
165 170 175	
GTC CCA TTT GTT CAG TGG TTC GTA GGG CTT TCC CCC ACT GTT TGG CTT	576
Val Pro Phe Val Gln Trp Phe Val Gly Leu Ser Pro Thr Val Trp Leu	
180 185 190	
TCA GTT ATA TGG ATG ATG TGG TAT TGG GGG CCA AGT CTG TAC AGC ATC	624
Ser Val Ile Trp Met Met Trp Tyr Trp Gly Pro Ser Leu Tyr Ser Ile	
195 200 205	
TTG AGT CCC TTT TTA CCG CTG TTA CCA ATT TTC TTT TGT CTT TGG GTA	672
Leu Ser Pro Phe Leu Pro Leu Leu Pro Ile Phe Phe Cys Leu Trp Val	
210 215 220	
TAC ATT TAA	681
Tyr Ile	
225	

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 226 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Glu Asn Ile Thr Ser Gly Phe Leu Gly Pro Leu Leu Val Leu Gln	
1 5 10 15	
Ala Gly Phe Phe Leu Leu Thr Arg Ile Leu Thr Ile Pro Gln Ser Leu	
20 25 30	
Asp Ser Trp Trp Thr Ser Leu Asn Phe Leu Gly Gly Thr Thr Val Cys	
35 40 45	
Leu Gly Gln Asn Ser Gln Ser Pro Thr Ser Asn His Ser Pro Thr Ser	
50 55 60	
Cys Pro Pro Thr Cys Pro Gly Tyr Arg Trp Met Cys Leu Arg Arg Phe	
65 70 75 80	
Ile Ile Phe Leu Phe Ile Leu Leu Leu Cys Leu Ile Phe Leu Leu Val	
85 90 95	
Leu Leu Asp Tyr Gln Gly Met Leu Pro Val Cys Pro Leu Ile Pro Gly	
100 105 110	
Ser Ser Thr Thr Ser Thr Gly Pro Cys Arg Thr Cys Met Thr Thr Ala	
115 120 125	
Gln Gly Thr Ser Met Tyr Pro Ser Cys Cys Cys Thr Lys Pro Ser Asp	
130 135 140	
Gly Asn Cys Thr Cys Ile Pro Ile Pro Ser Ser Trp Ala Phe Gly Lys	
145 150 155 160	

- 59 -

Phe Leu Trp Glu Trp Ala Ser Ala Pro Phe Ser Trp Leu Ser Leu Leu  
165 170 175

Val Pro Phe Val Gln Trp Phe Val Gly Leu Ser Pro Thr Val Trp Leu  
180 185 190

Ser Val Ile Trp Met Met Trp Tyr Trp Gly Pro Ser Leu Tyr Ser Ile  
195 200 205

Leu Ser Pro Phe Leu Pro Leu Leu Pro Ile Phe Phe Cys Leu Trp Val  
210 215 220

Tyr Ile  
225

- 60 -

What is claimed is:

CLAIMS

1. A method of inhibiting the replication of a naturally-occurring hepadnavirus, said method comprising the steps of:

5 (a) introducing into the proximity of said hepadnavirus a nucleic acid that encodes a hepadnavirus mutant polypeptide, wherein said polypeptide:

(i) comprises a first amino acid sequence that is substantially identical to a region of a wild  
10 type hepadnavirus core protein of at least 70 amino acids in length, and

(ii) lacks a second amino acid sequence of said wild type hepadnavirus core protein, wherein said second sequence comprises the carboxyterminal three amino  
15 acids of said wild type hepadnavirus core protein and does not exceed 100 amino acids in length; and

(b) allowing said mutant polypeptide to be expressed from said nucleic acid, wherein said mutant polypeptide inhibits the replication of said naturally-  
20 occurring hepadnavirus.

2. The method of claim 1, wherein said polypeptide further comprises a third amino acid sequence that is substantially identical to a portion of a wild type hepadnavirus surface protein.

25 3. The method of claim 1 or 2, wherein said hepadnavirus is hepatitis B virus (HBV).

4. The method of claim 1 or 2, wherein said method is used to treat an infection of hepatitis B virus in a patient.

- 61 -

5. The method of claim 3, wherein the carboxyterminal amino acid of said first amino acid sequence corresponds to a position selected from the group consisting positions 81 to 180 of SEQ ID NO: 12, inclusive.

6. The method of claim 3, wherein said carboxyterminal amino acid of said first amino acid sequence corresponds to a position selected from the group consisting of positions 171 to 180 of SEQ ID NO: 12, inclusive.

7. The method of claim 3, wherein said carboxyterminal amino acid of said first amino acid sequence corresponds to position 171 of SEQ ID NO: 12.

8. The method of claim 3, wherein said carboxyterminal amino acid of said first amino acid sequence corresponds to position 178 of SEQ ID NO: 12.

9. The method of claim 3, wherein said second amino acid sequence comprises amino acids 172-183 of SEQ ID NO: 12.

10. The method of claim 3, wherein the aminoterminal amino acid of said third amino acid sequence corresponds to a position selected from the group consisting of positions 1 to 112 of SEQ ID NO: 14, inclusive.

11. The method of claim 3, wherein the aminoterminal amino acid of said third amino acid sequence corresponds to a position selected from the group consisting of positions 1 to 8 of SEQ ID NO: 14, inclusive.

- 62 -

12. The method of claim 3, wherein said aminoterminal amino acid of said third amino acid sequence corresponds to position 5 of SEQ ID NO: 14.

13. The method of claim 3, wherein said  
5 aminoterminal amino acid of said third amino acid sequence corresponds to position 8 of SEQ ID NO: 14.

14. The method of claim 3, wherein the carboxyterminal amino acid of said third amino acid sequence corresponds to a position selected from the  
10 group consisting of positions 51 to 224 of SEQ ID NO: 14, inclusive.

15. The method of claim 3, wherein the carboxyterminal amino acid of said third amino acid sequence corresponds to a position selected from the  
15 group consisting of positions 112 to 224 of SEQ ID NO: 14, inclusive.

16. The method of claim 3, wherein the carboxyterminal amino acid of said third amino acid sequence corresponds to position 51 of SEQ ID NO: 14.

20 17. The method of claim 3, wherein the carboxyterminal amino acid of said third amino acid sequence corresponds to position 112 of SEQ ID NO: 14.

18. The method of claim 3, wherein the carboxyterminal amino acid of said third amino acid  
25 sequence corresponds to position 224 of SEQ ID NO: 14.

- 63 -

19. A method of inhibiting the replication of a naturally-occurring hepadnavirus, said method comprising introducing into the proximity of said hepadnavirus a hepadnavirus mutant polypeptide, wherein  
5 said polypeptide:

- (i) comprises a first amino acid sequence that is substantially identical to a region of a wild type hepadnavirus core protein of at least 70 amino acids in length, and
- 10 (ii) lacks a second amino acid sequence of said wild type hepadnavirus core protein, wherein said second sequence comprises the carboxyterminal three amino acids of said wild type hepadnavirus core protein and does not exceed 100 amino acids in length wherein said  
15 mutant polypeptide inhibits the replication of said hepadnavirus.

20. The method of claim 19, wherein said polypeptide further comprises a third amino acid sequence that is substantially identical to a portion of a wild  
20 type hepadnavirus surface protein.

- 64 -

21. A method of inhibiting the replication of a naturally-occurring hepadnavirus, said method comprising the steps of:

- (a) introducing into the proximity of said  
5 hepadnavirus a nucleic acid that encodes a hepadnavirus mutant polypeptide, wherein said polypeptide comprises:  
(i) a first amino acid sequence that is substantially identical to a region of a wild type hepadnavirus core protein of at least 70 amino acids, and  
10 (ii) a second amino acid sequence that is substantially identical to a portion of a wild type hepadnavirus surface protein.  
(b) allowing said mutant polypeptide to be expressed from said nucleic acid, wherein said mutant  
15 polypeptide inhibits the replication of said hepadnavirus.

22. A nucleic acid encoding a mutant hepatitis B virus (HBV) polypeptide, wherein said polypeptide:

- 20 (a) comprises a first amino acid sequence that is substantially identical to a region of a wild type HBV core protein of at least 70 amino acids in length; and  
(b) lacks a second amino acid sequence of said wild type HBV core protein, wherein said second sequence  
25 comprises the carboxyterminal three amino acids of said wild type HBV core protein and does not exceed nine amino acids in length.

23. The nucleic acid of claim 22, wherein the carboxyterminal amino acid of said first amino acid  
30 sequence is selected from the group consisting of the amino acids between position 174 and position 180 of SEQ ID NO: 12, inclusive.

- 65 -

24. A nucleic acid encoding a mutant hepadnavirus polypeptide, wherein said polypeptide:

(a) comprises a first amino acid sequence that is substantially identical to a region of a wild type  
5 hepadnavirus core protein of at least 70 amino acids in length;

(b) lacks a second amino acid sequence of said wild type hepadnavirus core protein, wherein said second sequence comprises the carboxyterminal three amino acids  
10 of said wild type hepadnavirus core protein; and

(c) comprises a third amino acid sequence that is substantially identical to a portion of a wild type hepadnavirus surface protein.

25. The nucleic acid of claim 24, wherein said  
15 second amino acid sequence does not exceed 100 amino acids in length.

26. The nucleic acid of claim 24, wherein the carboxyterminal amino acid of said first amino acid sequence corresponds to a position selected from the  
20 group consisting of positions 71 to 180 of SEQ ID NO: 12, inclusive.

27. A nucleic acid encoding a mutant hepadnavirus polypeptide, wherein said polypeptide comprises:

(a) a first amino acid sequence that is  
25 substantially identical to a region of a wild type hepadnavirus core protein of at least 70 amino acids in length; and

(b) a second amino acid sequence that is substantially identical to a portion of a wild type  
30 hepadnavirus surface protein.

- 66 -

28. A polypeptide encoded by the nucleic acid of any one of claims 22, 24, or 27.

29. A vector comprising the nucleic acid of any one of claims 22, 24, or 27.

5           30. A therapeutic composition comprising the mutant polypeptide of claim 28 in a pharmaceutically acceptable carrier.

10           31. A therapeutic composition comprising the vector of claim 29 in a pharmaceutically acceptable carrier.

15           32. The method of claim 1, wherein said hepadnavirus is selected from the group consisting of a woodchuck hepatitis virus (WHV), a hepatitis B virus (HBV), a hepatitis delta virus (HDV), a ground squirrel hepatitis B virus, and a duck hepatitis B virus (DHBV).

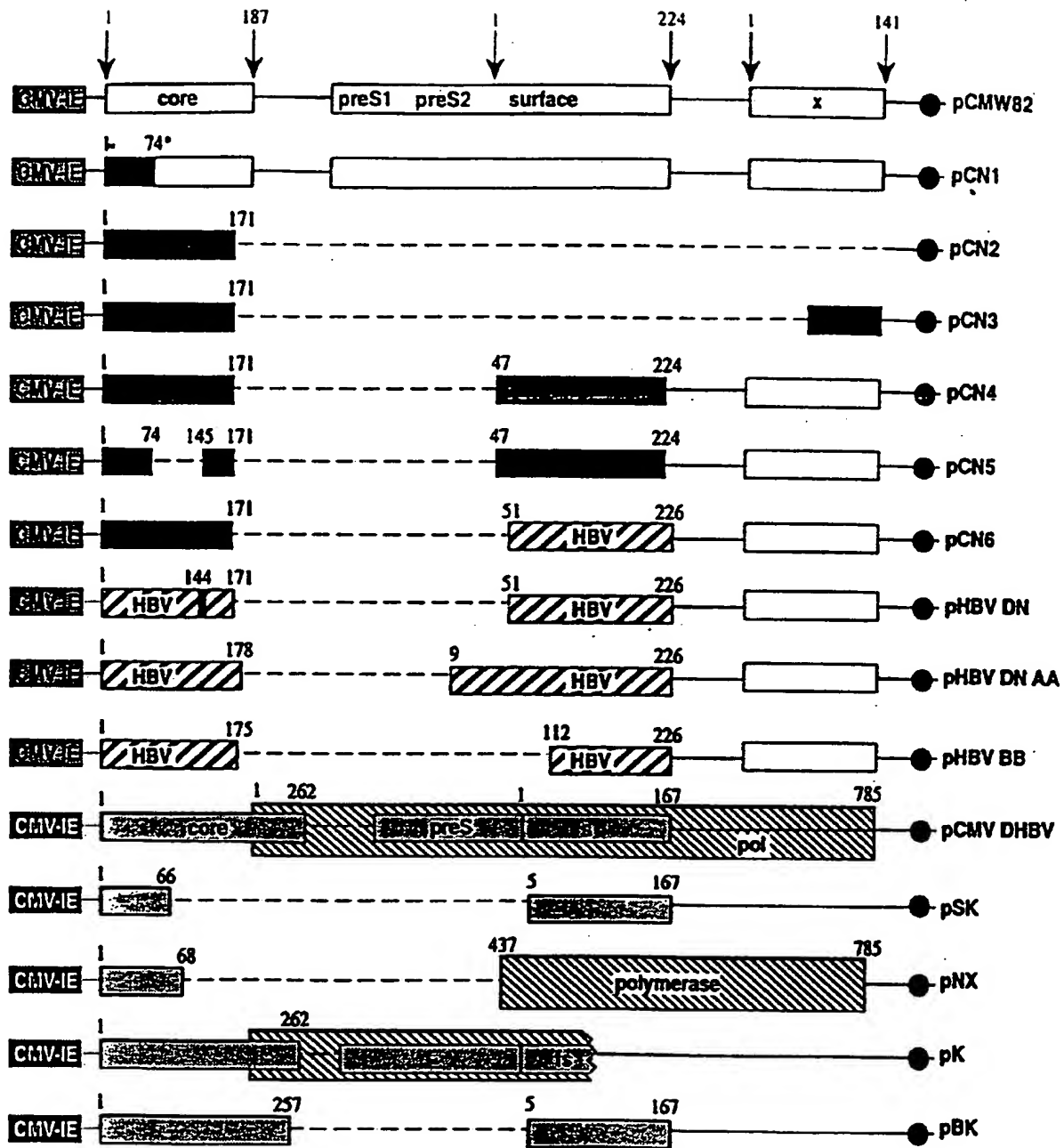
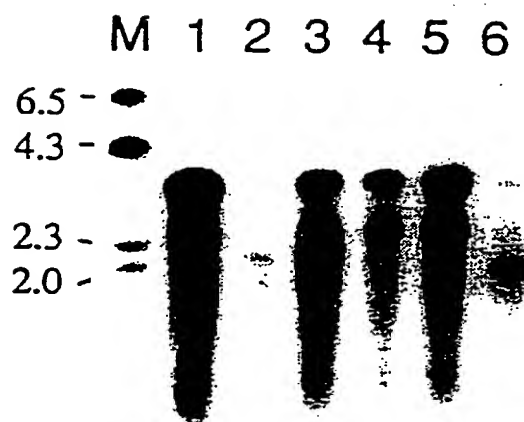
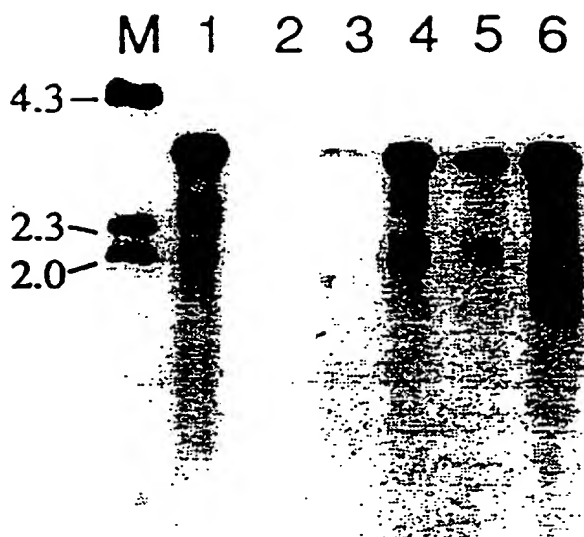


Fig 1

*Fig. 2*



*Fig. 3*



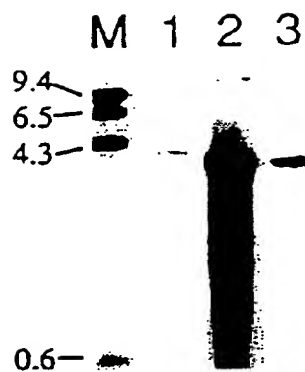


Fig. 4

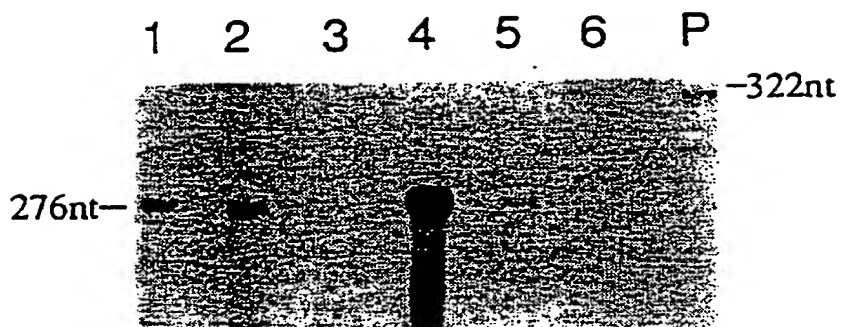
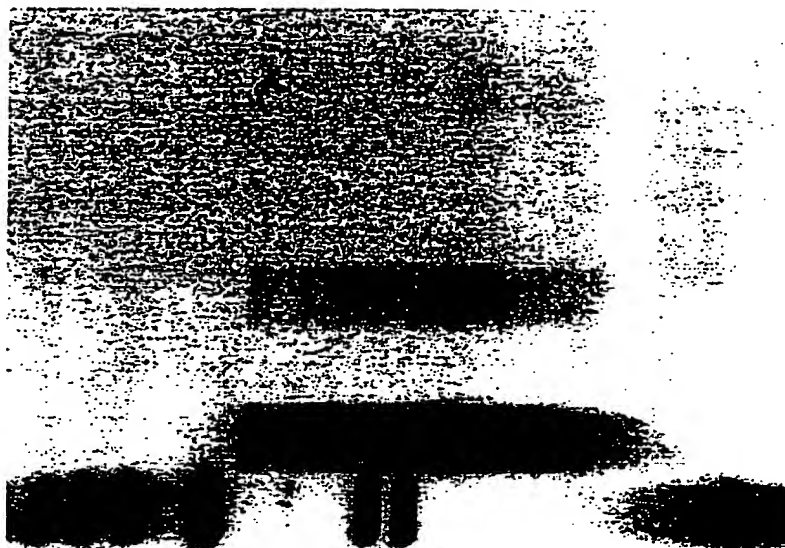


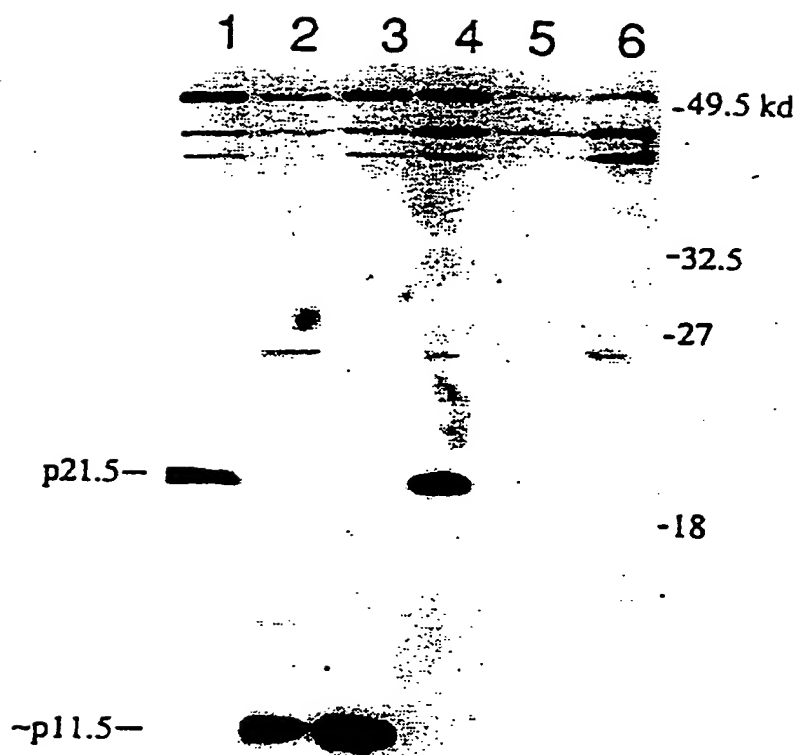
Fig. 6

Fig. 5

1234567



*Fig. 7*



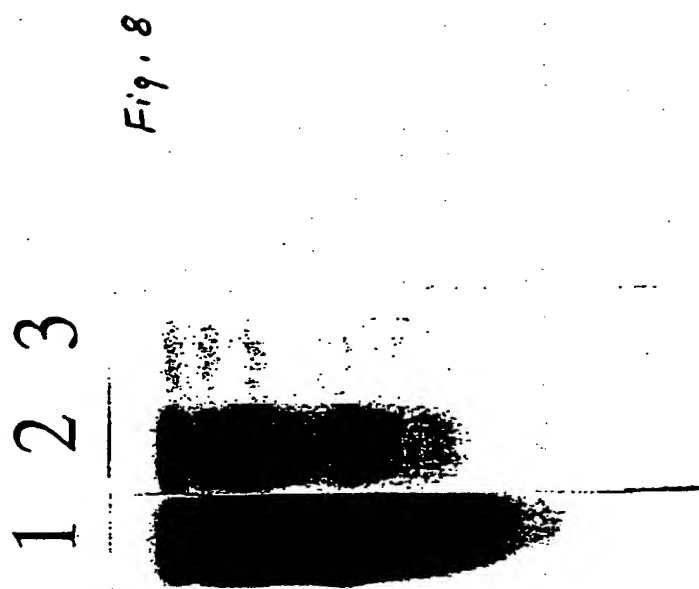


Fig. 9

1 2 3 4 5 6 7

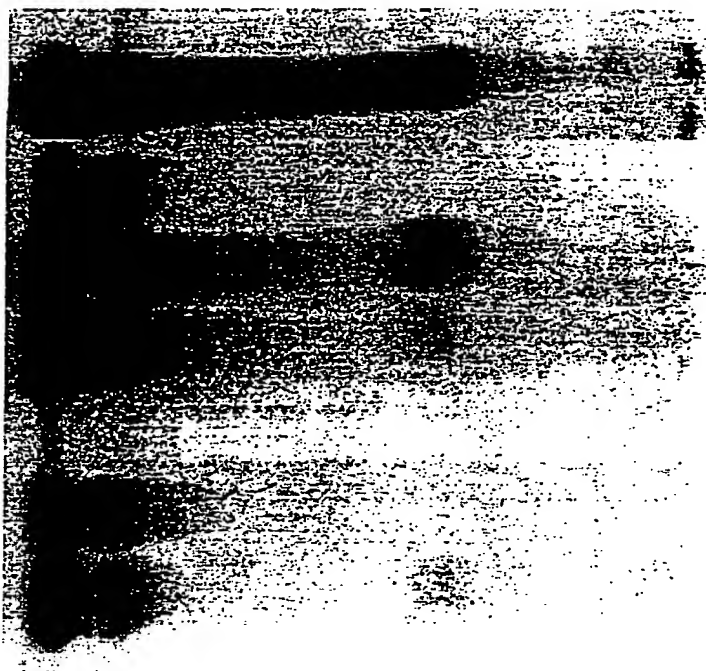


Fig. 10

```

1 / 1 31 / 11
ATG GAC ATA GAT CCC TAT AAA GAA TTT GGT TCA TCT TAT CAG TTG TTG AAT TTT CTT CCT
met asp ile asp pro tyr lys glu phe gly ser ser tyr gln leu leu asn phe leu pro
61 / 21 91 / 31
TTG GAC TTC TTT CCT GAC CTT AAT GCT TTG GTG GAC ACT GCT ACT GCC TTG TAT GAA GAA
leu asp phe phe pro asp leu asn ala leu val asp thr ala thr ala leu tyr glu glu
121 / 41 151 / 51
GAG CTA ACA GGT AGG GAA CAT TGC TCT CCG CAC CAT ACA GCT ATT AGA CAA GCT TTA GTA
glu leu thr gly arg glu his cys ser pro his his thr ala ile arg gln ala leu val
181 / 61 211 / 71
TGC TGG GAT GAA TTA ACT AAA TTG ATA GCT TGG ATG AGC TTT AAC ATA ACT TCT GAA CAA
cys trp asp glu leu thr lys leu ile ala trp met ser ser asn ile thr ser glu gln
241 / 81 271 / 91
GTA AGA ACA ATC ATA GTA AAT CAT GTC AAT GAT ACC TGG GGA CTT AAG GTG AGA CAA ACT
val arg thr ile ile val asn his val asn asp thr trp gly leu lys val arg gln ser
301 / 101 331 / 111
TTA TGG TTT CAT TTG TCA TOT CTC ACT TTC GGA CAA CAT ACA GTT CAA GAA TTT TTA GTA
leu trp phe his leu ser cys leu thr phe gly gln his thr val gln glu phe leu val
361 / 121 391 / 131
AGT TTT GTA GTA TGG ATC AGA ACT CCA GCT CCA TAT AGA CCT CCT AAT GCA CCC ATT CTC
ser phe val val trp ile arg thr pro ala pro tyr arg pro pro asn ala pro ile leu
421 / 141 451 / 151
TGC ACT CTT CCG GAA CAT ACA GTC ATT AGA AGA GGA GGT GCA AGA GCT TCT AGG TCC CCC
ser thr leu pro glu his thr val ile arg arg gly gly ala arg ala ser arg ser pro
481 / 161 511 / 171
AGA AGA CGC ACT CCC TCT CCT CGC AGG AGA AGA TCC CAA AAT TCG CAG TTC CAA ACT TGC
arg arg arg thr pro ser pro arg arg arg arg ser gln asn ser gln phe gln thr cys
541 / 181 571 / 191
AAA CAC TTG CCA ACC TCC TGT CCA CCA ACT TGC AAT GGC TTT CGT TGG ATG TAT CTG CCG
lys his leu pro thr ser cys pro pro thr cys asn gly phe arg trp met tyr leu arg
601 / 201 631 / 211
CGT TTT ATC ATA TAC CTA TTA GTC CTG CTG CTG TGC CTC ATC TTC TTG TTG GTT CTC CTG
arg phe ile ile tyr leu leu val leu leu leu cys leu ile phe leu leu val leu leu
661 / 221 691 / 231
GAC TGG AAA GGT TTA ATA CTT GTC TOT CTT CAA CCC ACA ACA GAA ACA ACA GTC AAT
asp trp lys gly leu ile pro val cys pro leu gln pro thr thr glu thr thr val asn
721 / 241 751 / 251
TGC AGA CAA TGC ACA ATC TCT GCA CAA GAC ATG TAT ACT CCT CCT TAC TGT TGT TGT TTA
cys arg gln cys thr ile ser ala gln asp met tyr thr pro pro tyr cys cys cys leu
781 / 261 811 / 271
AAA CCT ACG GCA GGA AAT TGC ACT TGT TGG CCC ATC CTT TCA TCA TGG GCT TTA GGA AAT
lys pro thr ala gly asn cys thr cys trp pro ile pro ser ser trp ala leu gly asn
841 / 281 871 / 291
TAC CTA TGG GAG TGG GCC TTA GCT CGT CTC TCT TGC CTC AAT TTA CTA GTG CCC TTG CTT
tyr leu trp glu trp ala leu ala arg leu ser trp leu asn leu leu val pro leu leu
901 / 301 931 / 311
CAA TGG TTA GGA GGA ATT TCC CTC ATT CCG TGG TTT TTG CTT ATA TGG ATG ATT TGG TTT
gln trp leu gly gly ile ser leu ile ala trp phe leu leu ile trp met ile trp phe
961 / 321 991 / 331
TGG GGG CCC GCA CTT CTG AGC ATC TTA CCG CCA TTT ATT CCC ATA TTT GTT CTG TTT TTC
trp gly pro ala leu leu ser ile leu pro pro phe ile pro ile phe val leu phe phe
1021 / 341
TTG ATT TGG GTA TAC ATT TGA
leu ile trp val tyr ile OPA

```

Fig. 11

```

1 / 1 31 / 11
ATG GAC ATC GAC CCT TAT AAA GAA TTT GGA GCT ACT GTG GAG TTA CTC TCG TTT TTG CCT
met asp ile asp pro tyr lys glu phe gly ala thr val glu leu leu ser phe leu pro
61 / 21 91 / 31
TCT GAC TTC TTT CCT TCA GTA CGA GAT CTT CTA GAT ACC GCC TCA GCT CTG TAT CGG GAA
ser asp phe phe pro ser val arg asp leu leu asp thr ala ser ala leu tyr arg glu
121 / 41 151 / 51
GCC TTA GAG TCT CCT GAG CAT TGT TCA CCT CAC CAT ACT GCA CTC AGG CAA GCA ATT CTT
ala leu glu ser pro glu his cys ser pro his his thr ala leu arg gln ala ile leu
181 / 61 211 / 71
TGC TGG GGG GAA CTA ATG ACT CTA GCT ACC TGG GTG GGT GTT AAT TTG GAA GAT CCA GCG
cys trp gly glu leu met thr leu ala thr trp val gly val asn leu glu asp pro ala
241 / 81 271 / 91
TCT AGA GAC CTA GTA GTC AGT TAT GTC AAC ACT AAT ATG GGC CTA AAG TTC AGG CAA CTC
ser arg asp leu val val ser tyr val asn thr asn met gly leu lys phe arg gln leu
301 / 101 331 / 111
TTG TGG TTT CAC ATT TCT TGT CTC ACT TTT GGA AGA GAA ACA GTT ATA GAG TAT TTG GTG
leu trp phe his ile ser cys leu thr phe gly arg glu thr val ile glu tyr leu val
361 / 121 391 / 131
TCT TTC GGA GTG TGG ATT CGC ACT CTT CCA GCT TAT AGA CCA CCA AAT GCC CCT ATC CTA
ser phe gly val trp ile arg thr pro pro ala tyr arg pro pro asn ala pro ile leu
421 / 141 451 / 151
TCA ACA CTT CCG GAA CAT ACA GTC ATT AGA AGA GGA GGT GCA AGA GCT TCT AGG TCC CCC
ser thr leu pro glu his thr val ile arg arg gly gly ala arg ala ser arg ser pro
481 / 161 511 / 171
AGA AGA CGC ACT CCC TCT CCT CGC AGG AGA AGA TCC CAA AAT TCG CAG TCC CCA ACC TCC
arg arg arg thr pro ser pro arg arg arg arg ser gln asn ser gln ser pro thr ser
541 / 181 571 / 191
AAT CAC TCA CCA ACC TCT TGT CCT CCA ACT TGT CCT GGT TAT CGC TGG ATG TGT CTG CGG
asn his ser pro thr ser cys pro pro thr cys pro gly tyr arg trp met cys leu arg
601 / 201 631 / 211
CGT TTT ATC ATC TTC CTC TTC ATC CTG CTG CTA TGC CTC ATC TTC TTG TTG GTT CTT CTG
arg phe ile ile phe leu phe ile leu leu leu cys leu ile phe leu leu val leu leu
661 / 221 691 / 231
GAC TAT CAX GGT ATG TTG CCC GTT TGT CCT CTA ATT CCA GGA TCC TCA ACA ACC AGC AGG
asp tyr gln gly met leu pro val cys pro leu ile pro gly ser ser thr thr ser thr
721 / 241 751 / 251
GGA CCA TGC CGG ACC TGC ATG ACT ACT GCT CAA GGA ACC TCT ATG TAT CCC TCC TGT TGC
gly pro cys arg thr cys met thr thr ala gln gly thr ser met tyr pro ser cys cys
781 / 261 811 / 271
TGT ACC AAA CCT TCG GAC CGA AAT TGC ACC TGT ATT CCC ATC CCA TCA TCC TGG GCT TTC
cys thr lys pro ser asp gly asn cys thr cys ile pro ile pro ser ser trp ala phe
841 / 281 871 / 291
GGA AAA TTC CTA TGG GAG TGG GCC TCA GCC CGT TTC TCC TGG CTC AGT TTA CTA GTC CCA
gly lys phe leu trp glu trp ala ser ala arg phe ser trp leu ser leu leu val pro
901 / 301 931 / 311
TTT GTT CAG TGG TTC GTA GGG CTT TCC CCC ACT GTT TGG CTT TCA GTT ATA TGG ATG ATG
phe val gln trp phe val gly leu ser pro thr val trp leu ser val ile trp met met
961 / 321 991 / 331
TGG TAT TGC CGC CCA ACT CTG TAC AGC ATC TTG AGT CCC TTT TTA CCG CTC TTA CCA ATT
trp tyr trp gly pro ser leu tyr ser ile leu ser pro phe leu pro leu leu pro ile
1021 / 341
TTC TTT TGT CTT TGG GTA TAC ATT TAA
phe phe cys leu trp val tyr ile OCH

```

Fig. 12

```

1 / 1 31 / 11
ATG GAC ATC GAG CCT TAT AAA GAA TTT GGA GCT ACT GTG GAG TTA CTC TCG TTT TTC CCT
met asp ile asp pro tyr lys glu phe gly ala thr val glu leu leu ser phe leu pro
61 / 21 91 / 31
TCT GAC TTC TTT CCT TCA GTA CGA GAT CTT CTA GAT ACC GCC TCA GCT CTG TAT CGG GAA
ser asp phe phe pro ser val arg asp leu leu asp thr ala ser ala leu tyr arg glu
121 / 41 151 / 51
GCC TTA GAG TCT CCT GAG CAT TGT TCA CCT CAC CAT ACT GCA CTC AGG CAA GCA ATT CTT
ala leu glu ser pro glu his cys ser pro his his thr ala leu arg gln ala ile leu
181 / 61 211 / 71
TGC TGG GGG GAA CTA ATG ACT CTA GCT ACC TGG GTG GGT GTT AAT TTC GAA GAT CCA GCG
cys trp gly glu leu met thr leu ala thr trp val gly val asn leu glu asp pro ala
241 / 81 271 / 91
TCT AGA GAC CTA GTA GTC AGT TAT GTC AAC ACT AAT ATG GCC CTA AAG TTC AGG CAA CTC
ser arg asp-leu val val ser tyr val asn thr asn met gly leu lys phe arg gln leu
301 / 101 331 / 111
TTG TGG TTT CAC ATT TCT TCT CTC ACT TTT GGA AGA GAA ACA GTT ATA GAG TAT TTG GTG
leu trp phe his ile ser cys leu thr phe gly arg glu thr val ile glu tyr leu val
361 / 121 391 / 131
TCT TTC GGA GTG TGG ATT CGC ACT CCT CCA GCT TAT AGA CCA CCA AAT GCC CCT ATC CTA
ser phe gly val trp ile arg thr pro pro ala tyr arg pro pro asn ala pro ile leu
421 / 141 451 / 151
TCA ACA CTT CCG GAG ACT ACT GTT GTT AGA CGA CGA GGC AGG TCC CCT AGA AGA AGA ACT
ser thr leu pro glu thr thr val val arg arg arg gly arg ser pro arg arg thr
481 / 161 511 / 171
CCC TCG CCT CGC AGA CGA AGG TCT CAA TCG CCG CGT CGC AGA AGA TCT CAA TCT CGC CTA
pro ser pro arg arg arg arg ser gln ser pro arg arg arg arg ser gln ser arg leu
541 / 181 571 / 191
GGA CCC CTT CTC GTG TTA CAG GCG GCG TTT TTC TTG TTA ACA AGA ATC CTC ACA ATA CCG
gly pro leu leu val leu gln ala gly phe phe leu leu thr arg ile leu thr ile pro
601 / 201 631 / 211
CAG AGT CTA GAC TCG TCG TCG ACT TCT CTC AAT TTT CTA GCG GGA ACT ACC GTG TGT CTT
gln ser leu asp ser trp trp thr ser leu asn phe leu gly gly thr thr val cys leu
661 / 221 691 / 231
CCC CAA AAT TCG CAG TCC CCA ACC TCC AAT CAC TCA CCA ACC TCT TGT CCT CCA ACT TGT
gly gln asn ser gln ser pro thr ser asn his ser pro thr ser cys pro pro thr cys
721 / 241 751 / 251
CCT GGT TAT CGC TCG ATG TGT CTG CCG CGT TTT ATC ATC TTC CTC TTC ATC CTG CTG CTA
pro gly tyr arg trp met cys leu arg arg phe ile ile phe leu phe ile leu leu leu
781 / 261 811 / 271
TGC CTC ATC TTC TTG TTG GTT CTT CTG GAC TAT CAA GGT ATG TTG CCC GTT TGT CCT CTA
cys leu ile phe leu leu val leu leu asp tyr gln gly met leu pro val cys pro leu
841 / 281 871 / 291
ATT CCA GGA TCC TCA ACA ACC AGC ACG GGA CCA TGC CCG ACC TGC ATG ACT ACT CCT CAA
ile pro gly ser ser thr thr ser thr gly pro cys arg thr cys met thr thr ala gln
901 / 301 931 / 311
GGA ACC TCT ATG TAT CCT TCC TGT TGC TGT ACC AAA CCT TCG GAC GGA AAT TGC ACC TGT
gly thr ser met tyr pro ser cys cys cys thr lys pro ser asp gly asn cys thr cys
961 / 321 991 / 331
ATT CCC ATC CCA TCA TCC TCG GCT TTC GGA AAA TTC CTA TGG GAG TGG GCC TCA GCC CGT
ile pro ile pro ser ser trp ala phe gly lys phe leu trp glu trp ala ser ala arg
1021 / 341 1051 / 351
TTC TCC TCG CTC AGT TTA CTA GTG CCA TTT GTT CAG TGG TTC GTA GCG CTT TCC CCC ACT
phe ser trp leu ser leu leu val pro phe val gln trp phe val gly leu ser pro thr
1081 / 361 1111 / 371
GTT TGG CTT TCA GTT ATA TGG ATG ATG TGC TAT TGG GCG CCA AGT CTG TAC AGC ATC TTG
val trp leu ser val ile trp met met trp tyr trp gly pro ser leu tyr ser ile leu
1141 / 381 1171 / 391
AGT CCC TTT TTA CCG CTG TTA CCA ATT TTC TTT TGT CTT TCG GTA TAC ATT TAA
ser pro phe leu pro leu leu pro ile phe phe cys leu trp val tyr ile OCH

```

## Fig. 13

```

1   /   1                               31  /   11
ATG GAC ATC GAC CCT TAT AAA GAA TTT GGA GCT ACT GTG GAG TTA CTC TCG TTT TTG CCT
met asp ile asp pro tyr lys glu phe gly ala thr val glu leu leu ser phe leu pro
61  /   21                               91  /   31
TCT GAC TTC TTT CCT TCA GTA CGA GAT CTT CTA GAT ACC GCC TCA GCT CTG TAT CGG GAA
ser asp phe phe pro ser val arg asp leu leu asp thr ala ser ala leu tyr arg glu
121 /   41                               151 /   51
GCC TTA GAG TCT CCT GAG CAT TGT TCA CCT CAC CAT ACT GCA CTC AGG CAA GCA ATT CTT
ala leu glu ser pro glu his cys ser pro his his thr ala leu arg gln ala ile leu
181 /   61                               211 /   71
TGC TGG GGG GAA CTA ATG ACT CTA GCT ACC TGG GTG GGT GTT AAT TTG GAA GAT CCA GCG
cys trp gly glu leu met thr leu ala thr trp val gly val asn leu glu asp pro ala
241 /   81                               271 /   91
TCT AGA GAC CTA GTA GTC AGT TAT GTC AAC ACT AAT ATC GGC CTA AAG TTC AGG CAA CTC
ser arg asp leu val val ser tyr val asn thr asn met gly leu lys phe arg gln leu
301 /  101                               331 /  111
TTG TGG TTT CAC ATT TCT TGT CTC ACT TTT GGA AGA GAA ACA GTT ATA GAG TAT TTG GTG
leu trp phe his ile ser cys leu thr phe gly arg glu thr val ile glu tyr leu val
361 /  121                               391 /  131
TCT TTC GGA GTG TGG ATT CGC ACT CCT CCA GCT TAT AGA CCA CCA AAT GCC CCT ATC CTA
ser phe gly val trp ile arg thr pro pro ala tyr arg pro pro asn ala pro ile leu
421 /  141                               451 /  151
TCA ACA CTT CCG GAG ACT ACT GTT GTT AGA CGA CGA GGC AGG TCC CCT AGA AGA AGA ACT
ser thr leu pro glu thr thr val val arg arg arg gly arg ser pro arg arg thr
481 /  161                               511 /  171
CCC TCG CCT CGC AGA CGA AGG TCT CAA TCG CCG CGT CGC AGA AGA TCG ATC CTC AAC AAC
pro ser pro arg arg arg arg ser gln ser pro arg arg arg arg ser ile leu asn asn
541 /  181                               571 /  191
CAG CAC GGG ACC ATG CCG GAC CTG CAT GAC TAC TGC TCA AGG AAC CTC TAT GTA TCC CTC
gln his gly thr met pro asp leu his asp tyr cys ser arg asn leu tyr val ser leu
601 /  201                               631 /  211
CTG TTG CTG TAC CAA ACC TTC GGA CGG AAA TTG CAC CTG TAT TCC CAT CCC ATC ATC CTG
leu leu leu tyr gln thr phe gly arg lys leu his leu tyr ser his pro ile ile leu
661 /  221                               691 /  231
GCC TTT CCG AAA ATT TCT ATG CGA GTG GGC CTC AGC CCG TTT CTC CTG GCT CAG TTT ACT
gly phe arg lys ile pro met gly val gly leu ser pro phe leu leu ala gln phe thr
721 /  241                               751 /  251
AGT GCC ATT TGT TCA GTG GTT CGT AGG GCT TTC CCC CAC TGT TTG GCT TTC AGT TAT ATC
ser ala ile cys ser val val arg arg ala phe pro his cys leu ala phe ser tyr met
781 /  261                               811 /  271
GAT GAT GTG GTA TTG GCG GCC AAG TCT GTA CAG CAT CTT GAG TCC CTT TTT ACC GCT GTT
asp asp val val leu gly ala lys ser val gln his leu glu ser leu phe thr ala val
841 /  281                               871 /  291
ACC AAT TTT CTT TTG TCT TTG GGT ATA CAT TTA
thr asn phe leu leu ser leu gly ile his leu

```

Fig. 14

1 / 1 31 / 11  
ATG GAT ATC AAT GCT TCT AGA GCC TTA GCC AAT GTG TAT GAT CTA CCA GAT GAT TTC TTT  
met asp ile asn ala ser arg ala leu ala asn val tyr asp leu pro asp asp phe phe  
61 / 21 91 / 31  
CCA AAA ATA GAT GAT CTT GTT AGA GAT GCT AAA GAC GCT TTA GAG CTT TAT TGG AAA TCA  
pro lys ile asp asp leu val arg asp ala lys asp ala leu glu pro tyr trp lys ser  
121 / 41 151 / 51  
GAT TCA ATA AAG AAA CAT GTT TTG ATT CCA ACT CAC TTT GTG GAT CTT ATT GAA GAC TTC  
asp ser ile lys lys his val leu ile ala thr his phe val asp leu ile glu asp phe  
181 / 61 211 / 71  
TGG CAG ACT ACA CAG GGC ATG CAT GAA ATA GCC GAA TCA TTA AGA GCT GTT ATA CCT CCC  
trp gln thr thr gln gly met his glu ile ala glu ser leu arg ala val ile pro pro  
241 / 81 271 / 91  
ACT ACT ACT CCT GTT CCA CCG GGT TAT CTT ATT CAG CAC GAA GAA GCT GAA GAG ATA CCT  
thr thr thr pro val pro pro gly tyr leu ile gln his glu glu ala glu glu ile pro  
301 / 101 331 / 111  
TTG GGA GAT TTA TTT AAA CAC CAA GAA GAA AGG ATA GTG AGT TTC CAA CCC GAC TAT CCG  
leu gly asp leu phe lys his gln glu glu arg ile val ser phe gln pro asp tyr pro  
361 / 121 391 / 131  
ATT ACG GCT AGA ATT CAT GCT CAT TTG AAA GCT TAT CCA AAA ATT AAC GAG GAA TCA CTG  
ile thr ala arg ile his ala his leu lys ala tyr ala lys ile asn glu glu ser leu  
421 / 141 451 / 151  
GAT ACG GCT ACG AGA TTG CTT TGG TGG CAT TAC AAC TGT TTA CTG TGG GGA GAA GCT CAA  
asp arg ala arg arg leu leu trp trp his tyr asn cys leu leu trp gly glu ala gln  
481 / 161 511 / 171  
GTT ACT AAC TAT ATT TCT CCG TTG CGT ACT TGG TTG TCA ACT CCT GAG AAA TAT AGA GGT  
val thr val val tyr ile ser arg leu arg thr trp leu ser thr pro glu lys tyr arg gly  
541 / 181 571 / 191  
AGA GAT GCC CCG ACC ATT GAA GCA ATC ACT AGA CCA ATC CAG GTG GCT CAG GGA GCC CGA  
arg asp ala pro thr ile glu ala ile thr arg pro ile gln val ala gln gly gly arg  
601 / 201 631 / 211  
AAA ACA ACT ACG GGT ACT AGA AAA CTT CGT GGA CTC GAA CCT AGA AGA AGA AAA GTT AAA  
lys thr thr thr gly thr arg lys pro arg gly leu glu pro arg arg arg lys val lys  
661 / 221 691 / 231  
ACC ACA GTT GTC TAT GGG AGA AGA CCG TCA AAG TCC CGG GCA ACG AGA GCC CCT ACA CCC  
thr thr val val tyr gly arg arg ser lys ser arg gly arg arg ala pro thr pro  
721 / 241 751 / 251  
CAA CGT CCG GCG TCC CCT CTC CCA CGT AGT TCG AGC AGC CAC CAT AGA TCC TTC GCG GGA  
gln arg ala gly ser pro leu pro arg ser ser ser ser his his arg ser phe gly gly  
781 / 261 811 / 271  
ATA CTA GCT GGC CTA ATC GGA TTA CTG GTA AGC TTT TTC TTG TTG ATA AAA ATT CTA GAA  
ile leu ala gly leu ile gly leu leu val ser phe phe leu leu ile lys ile leu glu  
841 / 281 871 / 291  
ATA CTG ACG AGG CTA GAT TGG TGG TGG ATT TCT CTC AGT TCT CCA AAG GGA AAA ATG CAA  
ile leu arg arg leu asp trp trp trp ile ser leu ser ser pro lys gly lys met gln  
901 / 301 931 / 311  
TGC GCT TTC CAA GAT ACT GGA GCC CAA ATC TCT CCA CAT TAC GTC GGA TCT TCC CCG TGG  
cys ala phe gln asp thr gly ala gln ile ser pro his tyr val gly ser cys pro trp  
961 / 321 991 / 331  
GGA TCC CCA GGA TTT CTT TGG ACC TAT CTC AGG CTT TTT ATC ATC TTC CTC TTA ATC CTG  
gly cys pro gly phe leu trp thr tyr leu arg leu phe ile ile phe leu leu ile leu  
1021 / 341 1051 / 351  
CTA GTA GCA GCA GGC TTG CTG TAT CTG ACG GAC AAC GGG TCT ACT ATT TTA GGA AAG CTC  
leu val ala ala gly leu leu tyr leu thr asp asn gly ser thr ile leu gly lys leu  
1081 / 361 1111 / 371  
CAA TGG GCG TCG GTC TCA GCC CTT TTC TCC TCC ATC TCT TCA CTA CTG CCC TCG GAT CCG  
gln trp ala ser val ser ala leu phe ser ser ile ser ser leu leu pro ser asp pro  
1141 / 381 1171 / 391  
AAA TCT CTC GTC GCT TTA ACG TTT GGA CTT TCA CTT ATA TGG ATG ACT TCC TCC TCT GCG  
lys ser leu val ala leu thr phe gly leu ser leu ile trp met thr ser ser ser ala  
1201 / 401 1231 / 411  
ACC CAA ACG CTC GTC ACC TTA ACG CAA TTA GCC ACG CTG TCT GCT CTT TTT TAC AAG AGC  
thr gln thr leu val thr leu thr gln leu ala thr leu ser ala leu phe tyr lys ser  
1261 / 421  
TAG

14/15

Fig. 15

[illegible]

Fig. 16

```

1 / 1 31 / 11
ATG CAG AAC ATC ACA TCA GGA TTC CTA GGA CCC CTT CTC GTG TTA CAG GCG GGG TTT TTC
met glu asn ile thr ser gly phe leu gly pro leu leu val leu gln ala gly phe phe
61 / 31
TTG TTC ACA AGA ATC CTC ACA ATA CCC CAC AGT CTA GAC TCG TGG TGG ACT TCT CTC AAT
leu leu thr arg ile leu thr ile pro gln ser leu asp ser trp trp thr ser leu asn
121 / 41
TTT CTA GCG GGA ACT ACC GTG TOT CTT GGC CAA AAT TCG CAG TCC CCA ACC TCC AAT CAC
phe leu gly gly thr thr val cys leu gly gln asn ser gln ser pro thr ser asn his
181 / 61
TCA CCA ACC TCT TGT CCT CCA ACT TGT CCT GGT TAT CGC TGG ATG TGT CTC CGG CGT TTT
ser pro thr ser cys pro pro thr cys pro gly tyr arg trp met cys leu arg arg phe
241 / 81
ATC ATC TTC CTC TTC ATC CTG CTG CTA TCC CTC ATC TTC TTG TTG GTT CTT CTC CAC TAT
ile ile phe leu phe ile leu leu leu cys leu ile phe leu leu val leu leu asp tyr
301 / 101
CAA GGT ATG TTG CCC GTT TGT CCT CTA ATT CCA GGA TCC TCA ACA ACC AGC ACC GGA CCA
gln gly met leu pro val cys pro leu ile pro gly ser ser thr thr ser thr gly pro
361 / 121
TCC CGC ACC TGC ATG ACT ACT GCT CAA CCA ACC TCT ATG TAT CCC TCC TGT TGC TGT ACC
cys arg thr cys met thr thr ala gln gly thr ser met tyr pro ser cys cys cys thr
421 / 141
AAA CCT TCG GAC GGA AAT TGC ACC TGT ATT CCC ATC CCA TCA TCC TGG GCT TTC GGA AAA
lys pro ser asp gly asn cys thr cys ile pro ile pro ser ser trp ala phe gly lys
481 / 161
TTC CTA TGG GAG TGG GCC TCA GCC CCT TTC TCC TGG CTC AGT TTA CTA CTC CCA TTT GTT
phe leu trp glu trp ala ser ala arg phe ser trp leu ser leu leu val pro phe val
541 / 181
CAG TCG TTC CTA GGG CTT TCC CCC ACT CTT TGG CTT TCA GTT ATA TGG ATC ATC TCG TAT
gln trp phe val gly leu ser pro thr val trp leu ser val ile trp met met trp tyr
601 / 201
TCC CGC CCA AGT CTG TAC AGC ATC TTG ACT CCC TTT TTA CCG CTG TTA CCA ATT TTC TTT
trp gly pro ser leu tyr ser ile leu ser pro phe leu pro leu leu pro ile phe phe
661 / 221
TGT CTT TGG GTA TAC ATT TAA
cys leu trp val tyr ile och

```

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US96/10602

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :A61K 48/00; C12N 5/00, 15/00  
US CL :435/172.3, 240.2, 320.1; 424/93.21; 514/44; 530/350  
According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/172.3, 240.2, 320.1; 424/93.21; 514/44; 530/350

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
APS MEDLINE BIOSIS CAPLUS WPIDS

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Journal of Virology, Volume 65, Number 04, issued April 1991, BLUM ET AL., "Naturally Occurring Missense Mutation in the Polymerase Gene Terminating Hepatitis B Virus Replication", pages 1836-1842, see entire document.	1-32
Y	Virology, Volume 199, issued 1994, MELEGARI ET AL., "Properties of Hepatitis B Virus Pre-S1 Deletion Mutants", pages 292-300, see entire document.	1-32

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*E* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*G* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search  
12 SEPTEMBER 1996

Date of mailing of the international search report  
30 SEP 1996

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Authorized officer

ANDREW MILNE

Facsimile No. (703) 305-3230

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)\*

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**